

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:30:15 ; Search time 21.5855 Seconds  
(without alignments)  
2286.817 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246

Sequence: 1 MSYKAAAGDYKADCPGPNP.....VTDEIKEMTPRKLSDFRQ 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US05\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US04\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US03\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/US02\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/US01\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US00\_NEW\_PUB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2246	100.0	437	9	US-09-813-718-12
2	2246	100.0	471	9	US-10-126-4678-2
3	2246	100.0	484	9	US-09-813-718-10
4	2231	99.3	471	9	US-09-919-039-163
5	2231	99.3	475	10	US-09-925-302-558
6	2116	94.2	415	9	US-09-813-718-14
7	1988	88.5	392	9	US-09-813-718-16
8	1129.5	50.3	433	9	US-10-128-714-8545
9	831	37.0	173	10	US-09-925-302-855
10	458	20.4	85	9	US-09-813-718-45
11	419	18.7	179	9	US-10-128-714-3545
12	398	17.7	85	9	US-09-813-718-46
13	363	16.2	85	9	US-09-813-718-48
14	357	15.9	85	9	US-09-813-718-47
15	356.5	15.9	142	10	US-09-925-302-557
16	256	11.4	46	9	US-09-813-718-51
17	210	9.3	46	9	US-09-813-718-52
18	185.5	8.3	341	10	US-09-815-242-13444
19	180	8.0	46	9	US-09-813-718-54

20	177	7.9	46	9	US-09-813-718-53	Sequence 53, Appl
21	147	6.5	385	9	US-10-128-714-3379	Sequence 3379, Ap
22	146.5	6.5	391	9	US-10-128-714-8379	Sequence 8379, Ap
23	146	6.5	339	10	US-09-815-242-11422	Sequence 11422, A
24	142.5	6.3	337	9	US-10-156-761-10954	Sequence 10954, A
25	134	6.0	372	9	US-09-813-718-8	Sequence 8, Appl1
26	132.5	5.9	372	9	US-09-813-718-4	Sequence 4, Appl1
27	132.5	5.9	536	9	US-09-813-718-2	Sequence 2, Appl1
28	128	5.7	334	10	US-09-815-242-11070	Sequence 11070, A
29	113.5	5.1	423	10	US-09-815-242-10678	Sequence 10678, A
30	113	5.0	415	10	US-09-815-242-5494	Sequence 5494, Ap
31	113	5.0	420	10	US-09-815-242-12649	Sequence 12649, A
32	112	5.0	345	9	US-09-738-626-4252	Sequence 4252, Ap
33	110.5	4.9	331	9	US-10-156-761-12259	Sequence 12259, A
34	108	4.8	306	9	US-10-126-927-50	Sequence 50, Appl
35	108	4.8	306	9	US-10-126-927-59	Sequence 59, Appl
36	108	4.8	306	9	US-10-126-927-64	Sequence 64, Appl
37	108	4.8	306	9	US-10-126-931A-50	Sequence 50, Appl
38	108	4.8	306	9	US-10-126-931A-59	Sequence 59, Appl
39	108	4.8	306	9	US-10-126-931A-64	Sequence 64, Appl
40	108	4.8	334	10	US-09-815-242-13806	Sequence 13806, A
41	107	4.8	306	9	US-10-126-927-51	Sequence 51, Appl
42	107	4.8	306	9	US-10-126-931A-51	Sequence 51, Appl
43	106.5	4.7	334	10	US-09-815-242-10371	Sequence 10371, A
44	106	4.7	306	9	US-10-126-927-45	Sequence 45, Appl
45	106	4.7	306	9	US-10-126-931A-45	Sequence 45, Appl

#### ALIGNMENTS

RESULT 1	US-09-813-718-12	US-09-813-718-12
Sequence 12, Appl1	Application US/09813718	
Publication No. US2002018266A1		
GENERAL INFORMATION:		
APPLICANT: Schimmel, Paul		
APPLICANT: Makasugi, Keisuke		
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For		
TITLE OF INVENTION: The Regulation of Angiogenesis		
FILE REFERENCE: 00-221		
CURRENT APPLICATION NUMBER: US/09/813,718		
CURRENT FILING DATE: 2001-03-21		
NUMBER OF SEQ ID NOS: 58		
SOFTWARE: PatentIn Ver. 2.0		
SEQ ID NO 12		
LENGTH: 437		
TYPE: PRT		
ORGANISM: Artificial Sequence		
FEATURE:		
OTHER INFORMATION: Description of Artificial Sequence: human mini		
OTHER INFORMATION: TPRS in PET20B		
US-09-813-718-12		
Query Match	100.0%	Score 2246; DB 9; Length 437;
Best Local Similarity	100.0%	Pred. No. 2.1e-203;
Matches 424; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
1 MSYKAAAGDYKADCPGPNPAPTSNHPGDATEABEDFVDPWTQTSAGKIDYDKLIVRF 60		
1 MSYKAAAGDYKADCPGPNPAPTSNHPGDATEABEDFVDPWTQTSAGKIDYDKLIVRF 60		
61 GSSKIDKELINIERATGPRPHFRGIFFSHRMNOVLDAVENKPPYLTTGGPSSSE 120		
61 GSSKIDKELINIERATGPRPHFRGIFFSHRMNOVLDAVENKPPYLTTGGPSSSE 120		
61 GSSKIDKELINIERATGPRPHFRGIFFSHRMNOVLDAVENKPPYLTTGGPSSSE 120		
61 GSSKIDKELINIERATGPRPHFRGIFFSHRMNOVLDAVENKPPYLTTGGPSSSE 120		
121 AAHVGHILPFIITKMLQDVNPLVLTQMTDDKCYLMKDLTDQAYGDAVENAKDIIACGF 180		
121 AAHVGHILPFIITKMLQDVNPLVLTQMTDDKCYLMKDLTDQAYGDAVENAKDIIACGF 180		
121 AAHVGHILPFIITKMLQDVNPLVLTQMTDDKCYLMKDLTDQAYGDAVENAKDIIACGF 180		
121 AAHVGHILPFIITKMLQDVNPLVLTQMTDDKCYLMKDLTDQAYGDAVENAKDIIACGF 180		
181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNVOKSIGFTSDSGTIGKISFPAIOAAP 240		
181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNVOKSIGFTSDSGTIGKISFPAIOAAP 240		
181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNVOKSIGFTSDSGTIGKISFPAIOAAP 240		
181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNVOKSIGFTSDSGTIGKISFPAIOAAP 240		

```

QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 300
DB 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 300
QY 301 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 360
DB 301 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 360
QY 361 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 420
DB 361 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 420
QY 421 FDFQ 424
DB 421 FDFQ 424

```

## RESULT 2

```

US-10-126-467B-2
; Sequence 2, Application US/10126467B
; Publication No. US2003005979A1
; GENERAL INFORMATION:
; APPLICANT: Paley, Elena
; TITLE OF INVENTION: ANIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE
; FILE REFERENCE: PALL-111
; CURRENT APPLICATION NUMBER: US/10/126,467B
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,980
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 09/513,895
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 09/384,869
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-126-467B-2

```

```

Query Match 100.0%; Score 2246; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 2.3e-203;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGKIDYDKLIVRF 60
DB 48 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGKIDYDKLIVRF 107
QY 61 GSSKIDKELINRERATGQRPHEFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 120
DB 108 GSSKIDKELINRERATGQRPHEFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 167
QY 121 AMVGHILIPFIETKMLQDVFNPLVIOMTDEKYLKMDLTLDQAYGAVENAKDIIACGF 180
DB 168 AMVGHILIPFIETKMLQDVFNPLVIOMTDEKYLKMDLTLDQAYGAVENAKDIIACGF 227
QY 181 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQYKGIFFGTDSDCIKISFPAIOAAP 240
DB 228 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQYKGIFFGTDSDCIKISFPAIOAAP 287
QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 300
DB 288 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 347
QY 301 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 360
DB 348 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 407
QY 361 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 420
DB 408 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 467

```

```

QY 421 FDFQ 424
DB 468 FDFQ 471

```

## RESULT 3

```

US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US2002018266A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: The Regulation of Angiogenesis
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human
US-09-813-718-10

```

```

Query Match 100.0%; Score 2246; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.4e-203;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGKIDYDKLIVRF 60
DB 48 MSYKAAAGEDYKADCPGNPAPTSNHPDTEAEEDFVDPWTVQTSAGKIDYDKLIVRF 107
QY 61 GSSKIDKELINRERATGQRPHEFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 120
DB 108 GSSKIDKELINRERATGQRPHEFLRGIFFSHRDNNQVLDAYENKKPFYLYTGRGSSSE 167
QY 121 AMVGHILIPFIETKMLQDVFNPLVIOMTDEKYLKMDLTLDQAYGAVENAKDIIACGF 180
DB 168 AMVGHILIPFIETKMLQDVFNPLVIOMTDEKYLKMDLTLDQAYGAVENAKDIIACGF 227
QY 181 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQYKGIFFGTDSDCIKISFPAIOAAP 240
DB 228 DINKTIFSDLDYMGSSGFYKVVVKIQKHVTFNQYKGIFFGTDSDCIKISFPAIOAAP 287
QY 241 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 300
DB 288 SFSNSFPQIRDRDIDICLIPCAIDODPYFRMTDVAIRIGYKPKALLHSTFPALOGAQ 347
QY 301 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 360
DB 348 TKMSASDPNSSIFLITDTAKOIKTKVKNKHAFCGSDTIEEHRQFGNCDVVSFMYLTFFL 407
QY 361 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 420
DB 408 EDDDKLEQIRKDYTSAGMLTGEIKKALIEVLQPLIAEHQARRKEVTDEIVKEFWTPRKL 467
QY 421 FDFQ 424
DB 468 FDFQ 471

```

## RESULT 4

```

US-09-919-039-163
; Sequence 163, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaefer, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

```

FILE REFERENCE: PA-0035 US  
CURRENT APPLICATION NUMBER: US/09/919,039  
CURRENT FILING DATE: 2002-09-09  
PRIOR APPLICATION NUMBER: 60/222,113  
PRIOR FILING DATE: 2000-07-28  
NUMBER OF SEQ ID NOS: 401  
SOFTWARE: PERL Program  
SEQ ID NO 163  
LENGTH: 471  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030108871A1 2705515CD1  
US-09-919-039-163

Query Match 99.3%; Score 2231; DB 9; Length 471;  
Best Local Similarity 99.5%; Pred. No. 6.1e-202;  
Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQSSAKGIDYDKLYRF 60  
DB 48 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQSSAKGIDYDKLYRF 107  
QY 61 GSSKIDKELINIRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 120  
DB 108 GSSKIDKELINIRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 167  
QY 121 AMHVGHLIPFTTKMLQDVFNVPVLIQWTDDEKYLMLKDLTDQAYGDAVENAKDIIACGF 180  
DB 168 AMHVGHLIPFTTKMLQDVFNVPVLIQWTDDEKYLMLKDLTDQAYGDAVENAKDIIACGF 227  
QY 181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIGKISFPALQAP 240  
DB 228 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIGKISFPALQAP 287  
QY 241 SFSNSFPQIFRDRTDIQCLIPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 300  
DB 288 SFSNSFPQIFRDRTDIQCLIPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 347  
QY 301 TMSASDPNSSIFLTDYKQIKTKVKNKHAFIGSGRDTIEHRQFGNCVDVDSFMYLTFFL 360  
DB 348 TMSASDPNSSIFLTDYKQIKTKVKNKHAFIGSGRDTIEHRQFGNCVDVDSFMYLTFFL 407  
QY 361 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKEVTDEIVKEFMTPRKLS 420  
DB 408 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKEVTDEIVKEFMTPRKLS 467  
QY 421 PDPQ 424  
DB 468 PDPQ 471

RESULT 5  
US-09-925-302-558  
Sequence 558, Application US/09925302  
Patent No. US20020044941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925,302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 558  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapiens

US-09-925-302-558

Query Match 99.3%; Score 2231; DB 10; Length 475;  
Best Local Similarity 99.5%; Pred. No. 6.2e-202;  
Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQSSAKGIDYDKLYRF 60  
DB 52 MSYKAAAGEDYKADCPGPNPAPTSNHPDTEAEEDFVDPMTVQSSAKGIDYDKLYRF 111  
QY 61 GSSKIDKELINIRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 120  
DB 112 GSSKIDKELINIRERATGQRPHHFLRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSE 171  
QY 121 AMHVGHLIPFTTKMLQDVFNVPVLIQWTDDEKYLMLKDLTDQAYGDAVENAKDIIACGF 180  
DB 172 AMHVGHLIPFTTKMLQDVFNVPVLIQWTDDEKYLMLKDLTDQAYGDAVENAKDIIACGF 231  
QY 181 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIGKISFPALQAP 240  
DB 232 DINKTFIFSDLDYMGSSGFYKNVVKIQKHVTFNQVKGIFGFTSDCIGKISFPALQAP 291  
QY 241 SFSNSFPQIFRDRTDIQCLIPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 300  
DB 292 SFSNSFPQIFRDRTDIQCLIPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 351  
QY 301 TMSASDPNSSIFLTDYKQIKTKVKNKHAFIGSGRDTIEHRQFGNCVDVDSFMYLTFFL 360  
DB 352 TMSASDPNSSIFLTDYKQIKTKVKNKHAFIGSGRDTIEHRQFGNCVDVDSFMYLTFFL 411  
QY 361 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKEVTDEIVKEFMTPRKLS 420  
DB 412 EDDDKLEQIRKDYTGAMLTGELKKALIEVLOPLAEHQARRKEVTDEIVKEFMTPRKLS 471  
QY 421 PDPQ 424  
DB 472 PDPQ 475

RESULT 6  
US-09-813-718-14  
Sequence 14, Application US/09813718  
Publication No. US2002018266A1  
GENERAL INFORMATION:  
APPLICANT: Schimmel, Paul  
APPLICANT: Wakasugi, Katsuke  
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
FILE REFERENCE: 00-221  
CURRENT APPLICATION NUMBER: US/09/813,718  
CURRENT FILING DATE: 2001-03-21  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 415  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human  
supermini typers in pet20B  
US-09-813-718-14

Query Match 94.2%; Score 2116; DB 9; Length 415;  
Best Local Similarity 100.0%; Pred. No. 3.6e-191;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SNHGDATEAEEDFVDPMTVQSSAKGIDYDKLYRFSSKIDKELINIRERATGQRPHH 83  
DB 2 SNHGDATEAEEDFVDPMTVQSSAKGIDYDKLYRFSSKIDKELINIRERATGQRPHH 61  
QY 84 FLRRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSEAMHVGHLIPFTTKMLQDVFNVP 143  
DB 62 FLRRGIFFSHRDMNVLDAYENKKPFYLYTGRGSSSEAMHVGHLIPFTTKMLQDVFNVP 121

```

QY 144 LV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPDINKTFIFSLDIDYMGSSGYKN 203
DB 122 LV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPDINKTFIFSLDIDYMGSSGYKN 181
QY 204 VK1QKAVTNOVGIFGFTSDICIGKISFPALQAASFSNFPQIRDRDIOCLIPCA 263
DB 182 VK1QKAVTNOVGIFGFTSDICIGKISFPALQAASFSNFPQIRDRDIOCLIPCA 241
QY 264 IDQPYFRMRDVAIRIGYPRKALLHSTFPALQAOTKMSASDPNSSIFLTDIAKQIKT 323
DB 242 IDQPYFRMRDVAIRIGYPRKALLHSTFPALQAOTKMSASDPNSSIFLTDIAKQIKT 301
QY 324 KVNHAASGGRTTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGEL 383
DB 302 KVNHAASGGRTTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGEL 361
QY 384 KKALIEVLQPLIAEHQARKEVTDEIVKEFWTPRKLSFDFQ 424
DB 362 KKALIEVLQPLIAEHQARKEVTDEIVKEFWTPRKLSFDFQ 402

```

## RESULT 7

```

US-09-813-718-16
; Sequence 16, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Antisense-cRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813, 718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 392
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human minor
; OTHER INFORMATION: Trps fragment in pr20B
US-09-813-718-16

```

Query Match 88.5%; Score 1988; DB 9; Length 392;  
Best Local Similarity 100.0%; Pred. No. 3,9e-179;

Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 47 SAKGIDVDKLVIRGSSKIDKELINRTERATGQRPHHFLRGIFFSHRDMNOVLDAYENK 106
DB 2 SAKGIDVDKLVIRGSSKIDKELINRTERATGQRPHHFLRGIFFSHRDMNOVLDAYENK 61
QY 107 KPFLYLRGSSSEAMVGHILPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYG 166
DB 62 KPFLYLRGSSSEAMVGHILPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYG 121
QY 167 DAVENAKDIIACGPDINKTFIFSLDIDYMGSSGYKNVVKIQKAVTNOVGIFGFTSD 226
DB 122 DAVENAKDIIACGPDINKTFIFSLDIDYMGSSGYKNVVKIQKAVTNOVGIFGFTSD 181
QY 227 CIGKISFPALQAASFSNFPQIRDRDIOCLIPCAIDQPYFRMRDVAIRIGYPRK 266
DB 182 CIGKISFPALQAASFSNFPQIRDRDIOCLIPCAIDQPYFRMRDVAIRIGYPRK 241
QY 287 LHSSTFPALQAOTKMSASDPNSSIFLTDIAKQIKTKVNKHAASGGRTTIEHRQFGN 346
DB 242 LHSSTFPALQAOTKMSASDPNSSIFLTDIAKQIKTKVNKHAASGGRTTIEHRQFGN 301
QY 347 CDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARKEVT 406
DB 302 CDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARKEVT 361

```

```

QY 407 DEIVKEFWTPRKLSFDFQ 424
DB 362 DEIVKEFWTPRKLSFDFQ 379

```

## RESULT 8

```

US-10-128-714-8545
; Sequence 8545, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengji
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Erosinkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285, 697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287, 066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295, 890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303, 899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316, 362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 8545
; LENGTH: 433
; TYPE: PRF
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8545

```

Query Match 50.3%; Score 1129.5; DB 9; Length 433;  
Best Local Similarity 53.7%; Pred. No. 4.4e-98;  
Matches 227; Conservative 64; Mismatches 115; Indels 17; Gaps 7;

```

QY 13 ADCEPBNAPT-----SNHGPDTEAEEDFVDPMVT-----OTSAGKIDYDKLIVRGSS 63
DB 2 ADCEPBNAPT-----SNHGPDTEAEEDFVDPMVT-----OTSAGKIDYDKLIVRGSS 60
QY 64 KIDKELINRTERATGQRPHHFLRGIFFSHRDMNOVLDAYENKPPFLYLRGSSSEAMH 123
DB 61 KIDKELINRTERATGQRPHHFLRGIFFSHRDMNOVLDAYENKPPFLYLRGSSSEAMH 120
QY 124 VGHILPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPD 182
DB 121 VGHILPFTKMLQDVNVNVLV1QMTDEKTLAKDLTDQAYGDAVENAKDIIACGPD 180
QY 183 NKTEIFSDLDYMGSSGYKNVVKIQKAVTNOVGIFGFTSDICIGKISFPALQAASPF 242
DB 181 NKTEIFSDLDYMGSSGYKNVVKIQKAVTNOVGIFGFTSDICIGKISFPALQAASPF 238
QY 243 SNFSPQIF-RDR---TDIOCLIPCAIDQPYFRMRDVAIRIGYPRKALLHSTFPALQ 298
DB 239 SNFSPQIF-RDR---TDIOCLIPCAIDQPYFRMRDVAIRIGYPRKALLHSTFPALQ 298
QY 299 AOTKMSASDPNSSIFLTDIAKQIKTKVNKHAASGGRTTIEHRQFGNCDVVSFMYLTF 358
DB 299 AOTKMSASDPNSSIFLTDIAKQIKTKVNKHAASGGRTTIEHRQFGNCDVVSFMYLTF 358
QY 359 FLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARKEVTDEIVKEFWTPRK 418
DB 359 FLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARKEVTDEIVKEFWTPRK 418
QY 419 LSF 421
DB 419 LSF 421

```





SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 46  
 LENGTH: 85  
 TYPE: PRT  
 ORGANISM: Bos taurus  
 US-09-813-718-46

Query Match  
 Best Local Similarity 88.1%; Score 398; DB 9; Length 85;  
 Matches 74; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 61  
 DB 2 SYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 61  
 DB 62 SSKIDKELINRIERATGQRPVHFL 85  
 DB 62 SSKIDKELINRIERATGQRPVHFL 85

## RESULT 13

US-09-813-718-48  
 Sequence 48; Application US/09813718  
 Publication No. US2002018266A1  
 GENERAL INFORMATION:  
 APPLICANT: Schimmel, Paul  
 TITLE OF INVENTION: Human Antinocetyl-cRNA Synthetase Polypeptides Useful For  
 TITLE OF INVENTION: The Regulation of Angiogenesis  
 FILE REFERENCE: 00-221  
 CURRENT APPLICATION NUMBER: US/09/813, 718  
 NUMBER OF SEQ ID NOS: 58  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 48  
 LENGTH: 85  
 TYPE: PRT  
 ORGANISM: Oryctolagus cuniculus  
 US-09-813-718-48

Query Match  
 Best Local Similarity 82.4%; Score 363; DB 9; Length 85;  
 Matches 70; Conservative 5; Mismatches 8; Indels 2; Gaps 2;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 60  
 DB 2 SYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 60  
 QY 61 GSSKIDKELINRIERATGQRPVHFL 85  
 DB 61 GSSKIDKELINRIERATGQRPVHFL 85

## RESULT 14

US-09-813-718-47  
 Sequence 47; Application US/09813718  
 Publication No. US2002018266A1  
 GENERAL INFORMATION:  
 APPLICANT: Schimmel, Paul  
 TITLE OF INVENTION: Human Antinocetyl-cRNA Synthetase Polypeptides Useful For  
 TITLE OF INVENTION: The Regulation of Angiogenesis  
 FILE REFERENCE: 00-221  
 CURRENT APPLICATION NUMBER: US/09/813, 718  
 NUMBER OF SEQ ID NOS: 58  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 47  
 LENGTH: 85  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-813-718-47

Query Match  
 Best Local Similarity 15.9%; Score 357; DB 9; Length 85;  
 Matches 69; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 60  
 DB 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 60  
 QY 61 GSSKIDKELINRIERATGQRPVHFL 85  
 DB 61 GSSKIDKELINRIERATGQRPVHFL 85

## RESULT 15

US-09-925-302-557  
 Sequence 557; Application US/09925302  
 Patent No. US20020044941A1  
 GENERAL INFORMATION:  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 FILE REFERENCE: PA104  
 CURRENT APPLICATION NUMBER: US/09/925,302  
 PRIOR FILING DATE: 2001-08-10  
 PRIOR APPLICATION NUMBER: PCT/US00/05918  
 PRIOR FILING DATE: 2000-03-08  
 PRIOR APPLICATION NUMBER: 60/124,270  
 NUMBER OF SEQ ID NOS: 896  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 557  
 LENGTH: 142  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (115)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (117)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (122)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (124)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (130)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (137)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 NAME/KEY: SITE  
 LOCATION: (142)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 US-09-925-302-557

Query Match  
 Best Local Similarity 15.9%; Score 356.5; DB 10; Length 142;  
 Matches 71; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 60  
 DB 52 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAEEDFVDPMTVQTSAGIDYDKLIVRF 111  
 QY 61 GSSKIDKELINRIERATGQRPVHFL 91  
 DB 112 GSSKIDKELINRIERATGQRPVHFL 139

Search completed: July 10, 2003, 12:33:36  
 Job time: 22.5855 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:24:54 ; Search time 33.92 Seconds

(without alignments)  
1665.633 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAAGBDYKADCPFGNP.....VTDEIVKFWTPKRLSPFQ 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2246	100.0	437	22	AA197616 Human mini TTPRS.
2	2246	100.0	437	23	AA197616 Human mini TTPRS.
3	2246	100.0	484	22	AA197615 Human full-length
4	2246	100.0	484	23	AA197615 Human full-length
5	2231	99.3	475	21	AA197615 Human tyrosophanyl
6	2226	99.1	471	20	AA197615 Human tyrosophanyl
7	2116	94.2	415	22	AA197617 Human HCW Inducib
8	2116	94.2	415	23	AA197617 Human HCW Inducib
9	1988	86.5	392	22	AA197618 Human supermini tr
10	1988	86.5	392	23	AA197618 Human supermini tr

11	1378	61.4	430	22	AB197616
12	1378	61.4	430	22	AB197616
13	1304	58.1	402	21	AA197616
14	1304	58.1	402	21	AA197616
15	1218.5	54.3	424	22	AA197616
16	968	43.1	292	21	AA197616
17	831	37.0	173	21	AA197616
18	803	35.8	365	22	AA197616
19	458	20.4	85	23	AA197616
20	356.5	15.9	142	21	AA197616
21	256	11.4	46	23	AA197616
22	203.5	8.3	341	19	AA197616
23	185.5	8.3	341	19	AA197616
24	185.5	8.3	341	19	AA197616
25	182.5	8.1	341	22	AA197616
26	174.5	7.8	341	23	AA197616
27	164.5	7.3	409	20	AA197616
28	162.5	7.2	409	20	AA197616
29	159.5	7.1	379	20	AA197616
30	158.5	7.1	379	20	AA197616
31	152	6.8	344	20	AA197616
32	146	6.5	339	22	AA197616
33	138.5	6.2	140	21	AA197616
34	134.5	6.0	525	22	AA197616
35	134	6.0	372	22	AA197616
36	134	6.0	372	22	AA197616
37	132.5	5.9	372	23	AA197616
38	132.5	5.9	372	23	AA197616
39	132.5	5.9	536	22	AA197616
40	132.5	5.9	536	22	AA197616
41	128	5.7	334	22	AA197616
42	126.5	5.6	419	23	AA197616
43	122.5	5.5	347	21	AA197616
44	122.5	5.5	346	21	AA197616
45	120.5	5.4	404	22	AA197616

# ALIGNMENTS

RESULT 1	AA197616	standard; Protein; 437 AA.
XX	AA197616	
XX	AA197616	
XX	07-JAN-2002 (first entry)	
XX	Human mini TTPRS.	
XX	Tyrosyl-tRNA synthetase; TyRS; Roessmann fold nucleotide binding domain;	
XX	vascular endothelial cell function; burn; plastic surgery; abdomen;	
XX	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;	
XX	dermal ulcer; diabetic ulcer; endotheialization;	
XX	tyrosophanyl-tRNA synthetase; tyRS; vascular graft surgery.	
OS	Homo sapiens.	
XX	WO200174841-A1.	
XX	11-OCT-2001.	
XX	21-MAR-2001; 2001WO-US08966.	
XX	31-MAR-2000; 2000US-193471P.	
XX	(SCRI) SCRIPS RES INST.	
XX	Schimmel P, Wakaugi K;	
XX	WPI; 2001-626377/72.	
XX	N-PSDB; AA197616.	
DR		

XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction  
XX  
XX  
XX Disclosure, Page 123-24; 150pp; English.

XX The sequences given in AAB47615-18 show full length and truncated  
CC versions of triptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kilo dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.

XX  
XX Sequence 437 AA;

Query Match 100.0%; Score 2246; DB 22; Length 437;  
Best Local Similarity 100.0%; Pred. No. 7, 4e-222;  
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYKAAAGEDYKADCPGNAPPTSNHGPDATEAEEDFVDEWTVQTSAAKIDYDKLIVRF 60  
DB 1 MSYKAAAGEDYKADCPGNAPPTSNHGPDATEAEEDFVDEWTVQTSAAKIDYDKLIVRF 60  
QY 61 GSSKIDKELINIRERATGCRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 120  
DB 61 GSSKIDKELINIRERATGCRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 120  
QY 121 AMHVGHLPIPTFKMLQDVNPLVIOMTDEKYLMDLTLDAYGDVAVENAKDIIACGF 180  
DB 121 AMHVGHLPIPTFKMLQDVNPLVIOMTDEKYLMDLTLDAYGDVAVENAKDIIACGF 180  
QY 121 AMHVGHLPIPTFKMLQDVNPLVIOMTDEKYLMDLTLDAYGDVAVENAKDIIACGF 180  
DB 121 AMHVGHLPIPTFKMLQDVNPLVIOMTDEKYLMDLTLDAYGDVAVENAKDIIACGF 180  
QY 181 DINKTIFSDLDVWGMSSGFYKVVVKIQKVTENQVGIFFGFTDSDCIKISFPALQAP 240  
DB 181 DINKTIFSDLDVWGMSSGFYKVVVKIQKVTENQVGIFFGFTDSDCIKISFPALQAP 240  
QY 241 SFSNSFPQIRDRDIOCLIPCAIDOPPRMRDVA PRIGYPPALHSTFPALQGAQ 300  
DB 241 SFSNSFPQIRDRDIOCLIPCAIDOPPRMRDVA PRIGYPPALHSTFPALQGAQ 300  
QY 241 SFSNSFPQIRDRDIOCLIPCAIDOPPRMRDVA PRIGYPPALHSTFPALQGAQ 300  
DB 241 SFSNSFPQIRDRDIOCLIPCAIDOPPRMRDVA PRIGYPPALHSTFPALQGAQ 300  
QY 301 TKMSASPNSSIFLTDAKQIKTKVNGHAFSGSDTIEHRQFGNCDVVSFWYLTFFL 360  
DB 301 TKMSASPNSSIFLTDAKQIKTKVNGHAFSGSDTIEHRQFGNCDVVSFWYLTFFL 360  
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKVETDEIYKEMTPPKLS 420  
DB 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKVETDEIYKEMTPPKLS 420  
QY 421 FDFQ 424  
DB 421 FDFQ 424

RESULT 2  
ID AAE13492  
XX AAE13492 standard; Protein; 437 AA.  
AC AAE13492;  
XX  
XX 12-FEB-2002 (first entry)

XX Human mini triptophanyl t-RNA synthetase in pET20B.  
XX  
XX Human; triptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;  
KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;  
KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
KW angiography; gene therapy; tumour; inflammation; vascular permeability;  
KW rheumatoid arthritis; psoriasis; diabetic retinopathy.  
XX  
XX Homo sapiens.  
XX  
XX WO200175078-A1.  
XX  
XX 11-OCT-2001.  
XX  
XX 21-MAR-2001; 2001WO-US08975.  
XX  
XX 31-MAR-2000; 2000US-193471P.  
XX  
XX (SCRI) SCRIPPS RES INST.  
XX  
XX Schimmel P, Wakaugl K;  
XX  
XX MPI; 2002-010784/01.  
XX N-PSDB; AAD22483.  
XX  
XX Novel truncated triptophanyl-tRNA synthetase polypeptides capable of  
PT regulating vascular endothelial cell function, preferably angiogenesis,  
PT is useful for treating solid tumor or suppressing tumor metastasis in  
PT mammal  
XX  
XX Example 1; Page 123-124; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly  
CC truncated triptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
CC fold nucleotide binding domain and polynucleotides encoding them. The  
CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
CC sequences are useful for regulating vascular endothelial cell function,  
CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
CC healing agents for re-vascularising damaged tissues. They are useful for  
CC treating full-thickness wounds (e.g. dermal ulcers), burns and injuries. TrpRS  
CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
CC sequences can also be used in plastic surgery when reconstruction is  
CC required following a burn, other trauma, or even for cosmetic purposes.  
CC Angiogenic TrpRS is also used in association with surgery and following  
CC the repair of cuts, for promoting endothelialisation in vascular graft  
CC surgery and for repairing the damage of myocardial infarction and in  
CC conjunction with coronary bypass surgery by stimulating the growth of  
CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
CC immunoassays to detect the presence of tumours. They are also useful  
CC for blocking endogenous angiogenic activity and retard the growth of  
CC solid tumours. These antibodies may also be used to treat inflammation  
CC caused by increased vascular permeability. Inhibiting the activity of  
CC TrpRS by antisense technology is useful for preventing further growth  
CC or even regress solid tumours, and for treating rheumatoid arthritis,  
CC psoriasis, diabetic retinopathy, all of which are characterised by  
CC abnormal angiogenesis. The present sequence is human truncated  
CC triptophanyl t-RNA synthetase (mini TrpRS; residues 48-471 of  
CC full-length TrpRS protein) protein in pET20B.

QY 1 MSYKAAAGEDYKADCPGNAPPTSNHGPDATEAEEDFVDEWTVQTSAAKIDYDKLIVRF 60  
DB 1 MSYKAAAGEDYKADCPGNAPPTSNHGPDATEAEEDFVDEWTVQTSAAKIDYDKLIVRF 60  
QY 61 GSSKIDKELINIRERATGCRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 120

```

Db 61 GSSKIDKELINRIERATGCPHFLRGIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 120
QY 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
Db 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
QY 181 DINKTFIFSDLDYMGSSGFYKVVVKIQKHTFNQVKGIFGFTSDCIGKISFPALQAP 240
Db 181 DINKTFIFSDLDYMGSSGFYKVVVKIQKHTFNQVKGIFGFTSDCIGKISFPALQAP 240
QY 241 SFSNSFPQIFRDRDTIQCLIPCAIDDPYFRMTDVAAPRIGYKPPALHSTFPALQAG 300
Db 241 SFSNSFPQIFRDRDTIQCLIPCAIDDPYFRMTDVAAPRIGYKPPALHSTFPALQAG 300
QY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKAFSGGRDTIEHRQFGNCDDVDSFMYLTFFL 360
Db 301 TKMSASDPNSSIFLDTAKQIKTKVKNKAFSGGRDTIEHRQFGNCDDVDSFMYLTFFL 360
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKVTDEIVKEMTPRKLS 420
Db 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKVTDEIVKEMTPRKLS 420
QY 421 PDFQ 424
Db 421 PDFQ 424

```

```

RESULT 3
AAB47615
ID AAB47615 standard; Protein; 484 AA.
XX
AC AAB47615;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human full-length TrpRS.
XX
KW Tyrosyl-tRNA synthetase; TrpRS; Roessmann fold nucleotide binding domain;
KW vascular endothelial cell function; burn; plastic surgery; abdomen;
KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW dermal ulcer; diabetic ulcer; endothelialization;
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.
XX
OS Homo sapiens.
XX
PN WO200174841-A1.
XX
PD 11-OCT-2001.
XX
PF 21-MAR-2001; 2001MO-US08966.
XX
PR 31-MAR-2000; 2000US-193471P.
XX
PA (SCRI ) SCRIIPS RRS INST.
XX
PI Schimmel P, Wakasugi K;
XX
DR WPI; 2001-626377/72.
XX
DR N-PSDB; AAA43602.
XX
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating
PT vascular endothelial function, in particular for regulating
PT angiogenesis, tumor metastasis and treating myocardial infarction
XX
PS Disclosure; Page 117-19; 150pp; English.
XX
CC The sequences given in AAB47615-18 show full length and truncated
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of
CC the invention comprises a Roessmann fold nucleotide binding domain, and
CC is capable of regulating vascular endothelial cell function. It is of
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of

```

CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic RNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.

Sequence 484 AA;

Query Match 100.0%; Score 2246; DB 22; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-222;  
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSYKAAAGDYADCPGPNPAPTSNHPDPAEEDFVDPMVTQSSAKGIDYDKLIYRF 60
Db 48 MSYKAAAGDYADCPGPNPAPTSNHPDPAEEDFVDPMVTQSSAKGIDYDKLIYRF 107
QY 61 GSSKIDKELINRIERATGCPHFLRGIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 120
Db 108 GSSKIDKELINRIERATGCPHFLRGIFFSHRDMNOVLDAVENKKPFYLTGRGPSS 167
QY 121 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 180
Db 168 AMHVGHLIPFIPTKMLQDVFNPLVIQMTDDEKYLKMDLTLDQAGDAVENAKDIIACGF 227
QY 181 DINKTFIFSDLDYMGSSGFYKVVVKIQKHTFNQVKGIFGFTSDCIGKISFPALQAP 240
Db 228 DINKTFIFSDLDYMGSSGFYKVVVKIQKHTFNQVKGIFGFTSDCIGKISFPALQAP 287
QY 241 SFSNSFPQIFRDRDTIQCLIPCAIDDPYFRMTDVAAPRIGYKPPALHSTFPALQAG 300
Db 288 SFSNSFPQIFRDRDTIQCLIPCAIDDPYFRMTDVAAPRIGYKPPALHSTFPALQAG 347
QY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKAFSGGRDTIEHRQFGNCDDVDSFMYLTFFL 360
Db 348 TKMSASDPNSSIFLDTAKQIKTKVKNKAFSGGRDTIEHRQFGNCDDVDSFMYLTFFL 407
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKVTDEIVKEMTPRKLS 420
Db 408 EDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKVTDEIVKEMTPRKLS 467
QY 421 PDFQ 424
Db 468 PDFQ 471

```

```

RESULT 4
AAE13491
ID AAE13491 standard; Protein; 484 AA.
XX
AC AAE13491;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human tryptophanyl t-RNA synthetase (TrpRS) in pET20B.
XX
KW Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;
KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;
KW re-vascularization; dermal ulcer; pressure sore; venous ulcer; injury;
KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;
KW angiography; gene therapy; tumour; inflammation; vascular permeability;
KW rheumatoid arthritis; psoriasis; diabetic retinopathy.
XX
OS Homo sapiens.

```

PN WO200175078-A1.  
 XX 11-OCT-2001.  
 PD 21-MAR-2001; 2001WO-US08975.  
 XX 31-MAR-2000; 2000US-193471P.  
 XX (SCRI ) SCRI PPS RES INST.  
 PA Schimmel P, Wakasugi K;  
 PI WPI; 2002-010784/01.  
 DR N-PSDB; AAD22482.  
 XX  
 XX Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
 PT regulating vascular endothelial cell function, preferably angiogenesis,  
 PT is useful for treating solid tumor or suppressing tumor metastasis in  
 PT mammal -  
 PS  
 PS Example 1; Page 117-119; 149pp; English.  
 XX  
 XX The patent discloses human aminoacyl tRNA synthetases, particularly  
 CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
 CC sequences are useful for regulating vascular endothelial cell function,  
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
 CC healing agents for re-vascularizing damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
 CC sequences can also be used in plastic surgery when reconstruction is  
 CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TrpRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
 CC immunoassays to detect the presence of tumors. They are also useful  
 CC for blocking endogenous angiogenic activity and retard the growth of  
 CC solid tumours. These antibodies may also be used to treat inflammation  
 CC caused by increased vascular permeability. Inhibiting the activity of  
 CC TrpRS by antisense technology is useful for preventing further growth  
 CC or even regress solid tumors, and for treating rheumatoid arthritis,  
 CC psoriasis, diabetic retinopathy, all of which are characterised by  
 CC abnormal angiogenesis. The present sequence is human tryptophanyl  
 CC t-RNA synthetase (TrpRS) in PET20B.  
 XX  
 XX  
 SQ Sequence 484 AA;  
 Query Match 100.0%; Score 2246; DB 23; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-222;  
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSYKAAAGEBYKADCPGPNPAPTSGNPGPDATAEEDFDPVTSAGKIDYKLVRF 60  
 DB 48 MSYKAAAGEBYKADCPGPNPAPTSGNPGPDATAEEDFDPVTSAGKIDYKLVRF 107  
 QY 61 GSSKIDKELINRRIRATGQRPHFLRRGIFPSHDMQVLDAYENKKPFYLYTGRGSSSE 120  
 DB 108 GSSKIDKELINRRIRATGQRPHFLRRGIFPSHDMQVLDAYENKKPFYLYTGRGSSSE 167  
 QY 121 AMHVGHILPIFTFTWLDVPRVPLVIQMTDEKYLMDLTLDQYGAVERNKKIICGF 180  
 DB 168 AMHVGHILPIFTFTWLDVPRVPLVIQMTDEKYLMDLTLDQYGAVERNKKIICGF 227  
 QY 181 DINKTFIFSDLDYMGSSGFYKYNVKIQKHTFNQVKGIFGFTSGDCIGKISFPAIDAP 240  
 DB 228 DINKTFIFSDLDYMGSSGFYKYNVKIQKHTFNQVKGIFGFTSGDCIGKISFPAIDAP 287  
 QY 241 SFSNSPQIFRDRITDIOCLIPCAIDODPYFPMRTDVAIRIGYPRALLHSTFPALQGAQ 300

DB 288 SFSNSPQIFRDRITDIOCLIPCAIDODPYFPMRTDVAIRIGYPRALLHSTFPALQGAQ 347  
 QY 301 TKMSASDPNNSITLTPTAKQIKTKNKAHPSGGRDITIEHRQSGCDVDVSMYLTFFL 360  
 DB 348 TKMSASDPNNSITLTPTAKQIKTKNKAHPSGGRDITIEHRQSGCDVDVSMYLTFFL 407  
 QY 361 EDDDKLEQIRKQDYSAMLTGELKKALIEVLQPLIAEHQARRKEVDEIVKEFMTPRKLS 420  
 DB 408 EDDDKLEQIRKQDYSAMLTGELKKALIEVLQPLIAEHQARRKEVDEIVKEFMTPRKLS 467  
 QY 421 FDFQ 424  
 DB 468 FDFQ 471  
 RESULT 5  
 AAB58220  
 ID AAB58220 standard; Protein; 475 AA.  
 XX  
 AC AAB58220;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polypeptide sequence SEQ ID 558.  
 XX  
 KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;  
 KW chondroactive; immunomodulatory; muscular active; vulnery;  
 KW gastrointestinal; nephrotoxic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055180-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05918.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Ruben SM;  
 XX  
 XX WPI; 2000-587514/55.  
 DR N-PSDB; AAF18096.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 PS Claim 11; Page 1052-1053; 1425pp; English.  
 XX  
 XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective, cytoskeletal, cardiovascular,  
 CC immunomodulatory, muscular active general, vulnery; gastrointestinal  
 CC general; nephrotoxic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer.  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.

XX SQ Sequence 475 AA;  
 Query Match 99.3%; Score 2231; DB 21; Length 475;  
 Best Local Similarity 99.3%; Pred. No. 2,9e-220;  
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPVWVQTSAGIDYDKLIYVF 60  
 52 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPVWVQTSAGIDYDKLIYVF 111  
 61 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 120  
 112 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 171

121 AMHVGLIPFIPTKMLQDVFNVLVYQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 180  
 172 AMHVGLIPFIPTKMLQDVFNVLVYQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 231  
 181 DINKTFIFSDLDYMGSSGFYKAVVVKIQKAVTFNQVKGIFGFTSDCIGKISFPALQAA 240  
 232 DINKTFIFSDLDYMGSSGFYKAVVVKIQKAVTFNQVKGIFGFTSDCIGKISFPALQAA 291  
 241 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFPALQAA 300  
 292 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFPALQAA 351  
 301 TKMSASDPNSSIFLTDTAQIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 360  
 352 TKMSASDPNSSIFLTDTAQIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 411  
 361 EDDDKLEQIRKDYTSGLMTELKALLIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 420  
 412 EDDDKLEQIRKDYTSGLMTELKALLIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 471

421 PDPQ 424  
 472 PDPQ 475

RESULT 6  
 AAY05372  
 ID AAY05372 standard; Protein; 471 AA.

AC AAY05372;  
 DT 30-JUN-1999 (first entry)  
 DE Human HCMV inducible gene protein, SEQ ID NO 12.  
 DE HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;  
 KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;  
 KM drug screening.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO913075-A2.  
 PD 18-MAR-1999.  
 PF 08-SEP-1998; 98WO-US18638.  
 PR 22-SEP-1997; 97US-0059725.  
 PR 08-SEP-1997; 97US-0058180.  
 PA (UYPF-) UNIV PRINCETON.  
 PI Cong J, Schenk T, Zhu H;  
 DR WPI, 1999-243729/20.  
 DR N-PSDB; AAX33942.  
 XX New isolated human genes  
 PT

XX PS Claim 3; Page 112-114; 184p; English.  
 XX CC This sequence is encoded by a human gene of the invention, and is induced  
 CC to express by both HCMV and interferon (IFN), designated HCMV-inducible  
 CC genes (c1g or c1gs). The invention also relates to genes that are  
 CC repressed in the presence of HCMV infection, designated HCMV-repressible  
 CC genes (crg or crgs). The products can be used to obtain agents which can  
 CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can  
 CC also be used for the development of drugs that would allow for higher  
 CC dosage IFN treatments without the concomitant toxicity normally  
 CC associated with administering high levels of IFN. The products can also  
 CC be used for detection, diagnosis and drug screening.

XX SQ Sequence 471 AA;  
 Query Match 99.1%; Score 2226; DB 20; Length 471;  
 Best Local Similarity 99.3%; Pred. No. 9.5e-220;  
 Matches 421; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPVWVQTSAGIDYDKLIYVF 60  
 48 MSYKAAAGEDYKADCPGPNAPTSNHPDPAEEDFVDPVWVQTSAGIDYDKLIYVF 107  
 61 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 120  
 108 GSSKIDKELINRIERATGQRPHHFLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSS 167

121 AMHVGLIPFIPTKMLQDVFNVLVYQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 180  
 168 AMHVGLIPFIPTKMLQDVFNVLVYQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 227  
 181 DINKTFIFSDLDYMGSSGFYKAVVVKIQKAVTFNQVKGIFGFTSDCIGKISFPALQAA 240  
 228 DINKTFIFSDLDYMGSSGFYKAVVVKIQKAVTFNQVKGIFGFTSDCIGKISFPALQAA 287  
 241 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFPALQAA 300  
 288 SFSNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFPALQAA 347  
 348 TKMSASDPNSSIFLTDTAQIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFL 407  
 361 EDDDKLEQIRKDYTSGLMTELKALLIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 420  
 408 EDDDKLEQIRKDYTSGLMTELKALLIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLS 467

421 PDPQ 424  
 468 PDPQ 471

RESULT 7  
 AAB47617  
 ID AAB47617 standard; Protein; 415 AA.

AC AAB47617;  
 DT 07-JAN-2002 (first entry)  
 DE Human supermini T19RS.  
 DE Tyrosyl-tRNA synthetase; T19RS; Rossmann fold nucleotide binding domain;  
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KM dermal ulcer; diabetic ulcer; endothelialization;  
 KM tryptophanyl-tRNA synthetase; t19RS; vascular graft surgery.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200174841-A1.

XX 11-OCT-2001.  
 PD 21-MAR-2001; 2001WO-US08966.  
 PF 31-MAR-2000; 2000US-193471P.  
 PR (SCRI ) SCRIIPS RES INST.  
 XX Schimmel P, Wakaugi K;  
 XX WPI; 2001-626377/72.  
 DR N-PSDB; AAA43604.  
 XX  
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating  
 PT angiogenesis, tumor metastasis and treating myocardial infarction -  
 XX  
 PS Disclosure; Page 129-30; 150pp; English.  
 XX  
 CC The sequences given in AAA47615-18 show full length and truncated  
 CC versions of tyrophanyl-tRNA synthetase (TyrRS). The truncated TyrRS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
 CC full length TyrRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TyrRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TyrRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.  
 CC  
 XX Sequence 415 AA;  
 SQ  
 Query Match 94.2%; Score 2116; DB 22; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-208;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8  
 AAEL13493  
 ID AAE13493 standard; Protein; 415 AA.  
 XX  
 AC AAE13493;  
 XX  
 DT 12-FEB-2002 (first entry)  
 XX  
 DE Human supermini tyrophanyl t-RNA synthetase in pET20B.  
 XX  
 KM Human; tyrophanyl-tRNA synthetase; TyrRS; tyrosyl t-RNA synthetase;  
 KM TyrRS; vascular endothelial cell function; angiogenesis; wound healing;  
 KM re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
 KM diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
 KM angiography; gene therapy; tumour; inflammation; vascular permeability;  
 KM rheumatoid arthritis; psoriasis; diabetic retinopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200175078-A1.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US08975.  
 XX  
 PR 31-MAR-2000; 2000US-193471P.  
 XX  
 PA (SCRI ) SCRIIPS RES INST.  
 XX  
 PI Schimmel P, Wakaugi K;  
 XX  
 DR WPI; 2002-010784/01.  
 DR N-PSDB; AAD22484.  
 XX  
 PT Novel truncated tyrophanyl-tRNA synthetase polypeptides capable of  
 PT regulating vascular endothelial cell function, preferably angiogenesis,  
 PT is useful for treating solid tumor or suppressing tumor metastasis in  
 PT mammal -  
 XX  
 PS Example 1; Page 129-130; 149pp; English.  
 XX  
 CC The patent discloses human aminoacyl tRNA synthetases, particularly  
 CC truncated tyrophanyl-tRNA synthetases (TyrRS) comprising a Rossmann  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TyrRS  
 CC sequences are useful for regulating vascular endothelial cell function,  
 CC preferably angiogenesis. Angiogenic TyrRS sequences are useful as wound  
 CC healing agents for re-vascularizing damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers), including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TyrRS  
 CC sequences can also be used in plastic surgery when reconstruction is  
 CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TyrRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TyrRS is also used in conjunction with angiography.  
 CC TyrRS DNAs are useful in gene therapy. TyrRS antibodies are used in  
 CC immunoassays to detect the presence of tumours. They are also useful  
 CC for blocking endogenous angiogenic activity and retard the growth of  
 CC solid tumours. These antibodies may also be used to treat inflammation  
 CC caused by increased vascular permeability. Inhibiting the activity of  
 CC TyrRS by antisense technology is useful for preventing further growth  
 CC or even regress solid tumours, and for treating rheumatoid arthritis,  
 CC psoriasis, diabetic retinopathy, all of which are characterised by  
 CC abnormal angiogenesis. The present sequence is human truncated  
 CC tyrophanyl t-RNA synthetase (supermini TyrRS; residues 71-471 of  
 CC full-length TyrRS protein) in pET20B.  
 XX  
 XX Sequence 415 AA;  
 SQ



Query Match 94.2%; Score 2116; DB 23; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-208;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SNEGPDAEEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPBH 83  
 DB 2 SNEGPDAEEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPBH 61  
 QY 84 FLRGGIFFSHRDMNOVLDAVENKKPFLYTGRRGSSAMHVGHLIPFTKWLQDVFNVP 143  
 DB 62 FLRGGIFFSHRDMNOVLDAVENKKPFLYTGRRGSSAMHVGHLIPFTKWLQDVFNVP 121  
 QY 144 LVIQMTDEKYLMDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
 DB 122 LVIQMTDEKYLMDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 QY 204 VKKIQKVTFNQYKGFPGTSDSCIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCA 263  
 DB 182 VKKIQKVTFNQYKGFPGTSDSCIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCA 241  
 QY 264 IDDDPYFRMTRDVAAPRIGYKPPALHSTFPPALQAGOTKMSASDPNSSIFLDTAKOIKT 323  
 DB 242 IDDDPYFRMTRDVAAPRIGYKPPALHSTFPPALQAGOTKMSASDPNSSIFLDTAKOIKT 301  
 QY 324 KVNKHA FSGGRDTIEHRQFGNCVDVSEMYLTFLEDDDKLEQIRKDYTGAMLTGEL 383  
 DB 302 KVNKHA FSGGRDTIEHRQFGNCVDVSEMYLTFLEDDDKLEQIRKDYTGAMLTGEL 361  
 QY 384 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFO 424  
 DB 362 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFO 402

RESULT 9  
 AAB47618  
 ID AAB47618 standard; Protein; 392 AA.  
 XX AAB47618;  
 XX 07-JAN-2002 (first entry)  
 DE Human inactive Tryps.  
 XX Tyrosyl-tRNA synthetase; Tryps; Roessmann fold nucleotide binding domain;  
 KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;  
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW tryptophanyl-tRNA synthetase; tryps; vascular graft surgery.  
 OS Homo sapiens.  
 XX  
 XX WO200174841-A1.  
 PN  
 XX  
 PD 11-OCT-2001.  
 PF 21-MAR-2001; 2001MO-US08966.  
 PR 31-MAR-2000; 2000US-193471P.  
 PA (SCRI ) SCRIPPS RES INST.  
 PA Schimmel P, Wakaugui K;  
 PI WPI, 2001-626377/72.  
 DR N-PSDB; AAH43605.  
 XX  
 XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating  
 PT angiogenesis, tumor metastasis and treating myocardial infarction -  
 XX Disclosure; Page 135-36; 150pp; English.  
 XX

The sequences given in AAB47618 show full length and truncated  
 CC versions of tryptophanyl-tRNA synthetase (Tryps). The truncated Tryps of  
 CC the invention comprises a Roessmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kilo dalton molecular weight and is produced by cleavage of  
 CC full length Tryps with polymorphonuclear leucocyte elastase. Truncated  
 CC Tryps is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated Tryps  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.  
 XX

SO Sequence 392 AA;  
 Query Match 88.5%; Score 1988; DB 22; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-195;  
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPBHFLRGGIFFSHRDMNOVLDAVENK 106  
 DB 2 SAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPBHFLRGGIFFSHRDMNOVLDAVENK 61  
 QY 107 KFFLYLTGRRGSSAMHVGHLIPFTKWLQDVFNVPVLIQMTDEKYLMDLTDQAYG 166  
 DB 62 KFFLYLTGRRGSSAMHVGHLIPFTKWLQDVFNVPVLIQMTDEKYLMDLTDQAYG 121  
 QY 167 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIQRKDYTGAMLTGEL 226  
 DB 122 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIQRKDYTGAMLTGEL 181  
 QY 227 CIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPA 286  
 DB 182 CIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDDDPYFRMTRDVAAPRIGYKPA 241  
 QY 287 LHSFPPALQAGOTKMSASDPNSSIFLDTAKOIKTKVNKHA FSGGRDTIEHRQFGN 346  
 DB 242 LHSFPPALQAGOTKMSASDPNSSIFLDTAKOIKTKVNKHA FSGGRDTIEHRQFGN 301  
 QY 347 CDVDVSEMYLTFLEDDDKLEQIRKDYTGAMLTGELKKAILEVLOPLIAEHQARRKEVT 406  
 DB 302 CDVDVSEMYLTFLEDDDKLEQIRKDYTGAMLTGELKKAILEVLOPLIAEHQARRKEVT 361  
 QY 407 DEIVKEFMTPRKLSFDFO 424  
 DB 362 DEIVKEFMTPRKLSFDFO 379

RESULT 10  
 AAE13494  
 ID AAE13494 standard; Protein; 392 AA.  
 XX AAE13494;  
 XX 12-FEB-2002 (first entry)  
 DE Human inactive tryptophanyl t-RNA synthetase in pET20B.  
 XX  
 XX Human; tryptophanyl-tRNA synthetase; Tryps; tyrosyl t-RNA synthetase;  
 KW Tryps; vascular endothelial cell function; angiogenesis; wound healing;  
 KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
 KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
 KW angiography; gene therapy; tumour; inflammation; vascular permeability;  
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy.  
 OS Homo sapiens.  
 XX

XX MO200175078-A1.  
 PN 11-OCT-2001.  
 XX 21-MAR-2001; 2001WO-US08975.  
 XX 31-MAR-2000; 2000US-193471P.  
 XX (SCRI ) SCRIIPS RES INST.  
 XX Schimmel P, Wakaugi K;  
 XX WPI; 2002-010784/01.  
 DR N-PSDB; AAD22485.  
 XX  
 PT Novel truncated cryptophanyl-tRNA synthetase polypeptides capable of  
 PT regulating vascular endothelial cell function, preferably angiogenesis,  
 PT is useful for treating solid tumor or suppressing tumor metastasis in  
 PT mammal  
 XX  
 XX Disclosure, Page 135-136, 149pp; English.  
 XX  
 XX The patent discloses human aminoacyl tRNA synthetases, particularly  
 CC truncated cryptophanyl-tRNA synthetases (TrpRS) comprising a Roseman  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
 CC sequences are useful for regulating vascular endothelial cell function,  
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
 CC healing agents for re-vascularizing damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
 CC sequences can also be used in plastic surgery when reconstruction is  
 CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TrpRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
 CC immunoassays to detect the presence of tumours. They are also useful  
 CC for blocking endogenous angiogenic activity and retard the growth of  
 CC solid tumours. These antibodies may also be used to treat inflammation  
 CC caused by increased vascular permeability. Inhibiting the activity of  
 CC TrpRS by antisense technology is useful for preventing further growth  
 CC or even regress solid tumours, and for treating rheumatoid arthritis,  
 CC psoriasis, diabetic retinopathy, all of which are characterised by  
 CC abnormal angiogenesis. The present sequence is human inactive  
 CC cryptophanyl t-RNA synthetase (TrpRS) in pET20B.  
 XX  
 XX Sequence 392 AA;  
 SQ  
 Query Match 88.5%; Score 1988; DB 23; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 2,2e-195;  
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 47 SAKSIDYDKLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENK 106  
 DB 2 SAKSIDYDKLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENK 61  
 QY 107 KPFLYLTGRGPSSSAMVGHILPITFTKWLQDVNVNVLVIMTDEKYLWMDLTDQAYG 166  
 DB 62 KPFLYLTGRGPSSSAMVGHILPITFTKWLQDVNVNVLVIMTDEKYLWMDLTDQAYG 121  
 QY 167 DAVENADIIACGDIKNTPIFSDLIDYMGSSGFGYKAVVVKIQKAVTNQVKGIFGFDSD 226  
 DB 122 DAVENADIIACGDIKNTPIFSDLIDYMGSSGFGYKAVVVKIQKAVTNQVKGIFGFDSD 181  
 QY 227 CIGGISPPAIQAAFSFNSPQIRDRDIDICLPCAIIDDPYFRMRDVAAPRGYKPA 286  
 DB 182 CIGGISPPAIQAAFSFNSPQIRDRDIDICLPCAIIDDPYFRMRDVAAPRGYKPA 241  
 QY 287 LHSSTFPALQAGQTKMSASDPNSSIFLTDYAKQIKTKVNGAFSGGRDTIEHRQFGGN 346

DB 242 LHSSTFPALQAGQTKMSASDPNSSIFLTDYAKQIKTKVNGAFSGGRDTIEHRQFGGN 301  
 QY 347 CDVDVSEFMVLTFFLEDDDKLEQIRKDYTSAMLTGELKALIEVLOPLIAEHQARKKEYT 406  
 DB 302 CDVDVSEFMVLTFFLEDDDKLEQIRKDYTSAMLTGELKALIEVLOPLIAEHQARKKEYT 361  
 QY 407 DEIVKEFMTPRKLSFDQ 424  
 DB 362 DEIVKEFMTPRKLSFDQ 379  
 RESULT 11  
 ABB64621  
 ID ABB64621 standard; Protein; 430 AA.  
 XX  
 AC ABB64621;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 20655.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN NO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PMD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL08724.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA  
 CC sequences (ABU1840-ABU16175) and the encoded proteins  
 CC (ABBS7737-ABBS7072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at fcp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 430 AA;  
 Query Match 61.4%; Score 1378; DB 22; Length 430;  
 Best Local Similarity 61.4%; Pred. No. 1.3e-132;  
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;  
 QY 10 DYKADCPGNGAPTSNHGPD-----TEAE-----EDFVDEMTVQTSAGKIDYD 54  
 DB 3 DTEKTEVEGEALTLNGKPDABEVEGTDAQAQEGATAPEDVVDPMNVASSDAGVDYD 62  
 QY 55 KLIVRGSSKIDKELINRIERATGQRPHPLRGIFSRDMNQVDAVENKPFYLYTG 114

DB 63 KLIRKRGSSKIDELIARFEKIKGKPAHHFIRRGMEFSSHRDLHTLITLREQKPFYLYTG 122  
 QY 115 RGPSSAMVGHILIPFIPTKWLQDVFNVPVLTQMTDEKYLWKDLTLDOAYGDAVENAKD 174  
 DB 123 RGPSSGSLHVGHLVFPIMTKWLQETFDVPLVQLTDEKTLMKDLKVEDAIKLGRENAD 182  
 QY 175 IIAAGFDINKTIFESDLDMYKMSGFGYKNNVKIQKHVTFNQVKGIFGFTSDICIGKISFP 234  
 DB 183 IYALGFDVKTFFIENNLFEVKGCPAMYQNIIRIQKCVTFNQVKGIFGFTSDICIGKISFP 242  
 QY 235 AIOAAPSFNSFPQIFRDRDTQCLIPCAIDDPYFRMTDVAIRIGYKPKALLHSTFEP 294  
 DB 243 AAQAAAPALISTPFPFIGNR-KVHCILPCALIDDPYFRMTDVAIRIGYKPKALLHSTFEP 301  
 QY 295 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRDTIEHRQPGNCDDVVSFM 354  
 DB 302 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRDTIEHRQPGNCDDVVSFM 361  
 QY 355 YLTFPLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLAEHQARRKREVTDEIVKEFM 414  
 DB 362 LKTFPLEDDDKLEEVAVVASKGEMLTGEIKGLAVETLTPVEHQAAARKLITDEVLDKTF 421  
 QY 415 TPRKLSF 421  
 DB 422 ELRPLKF 428

RESULT 12  
 ABB67203  
 ID ABB67203 standard; Protein; 430 AA.  
 AC ABB67203;  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 XX Drosophila melanogaster polypeptide SEQ ID NO 28401.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX  
 XX N-PSDB; ABL11306.  
 DR  
 DR WPI; 2001-656860/75.  
 XX  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX  
 PS Disclosure; SEQ ID NO 28401; 21bp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB11840-AB116175) and the encoded proteins  
 CC (ABBS7737-ABBS7072).  
 CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX  
 XX Sequence 430 AA;  
 QY  
 DB 10 DYKADCPGNPAPTSNHGPA-----TEAE-----EDFVDPWTQTSANGIDVD 54  
 DB 3 DTKETVEGVETLTLNGKEDAPVETGTDAAQOEGATAPTEDVDVPMVNAASNDAGVDYD 62  
 QY 55 KLIVRPGSSKIDKELINRTERATGQRPHHFLRAGIFPSHRDMNQVLDAVENKKPFYLYTG 114  
 DB 63 KLIRKRGSSKIDELIARFEKIKGKPAHHFIRRGMEFSSHRDLHTLITLREQKPFYLYTG 122  
 QY 115 RGPSSAMVGHILIPFIPTKWLQDVFNVPVLTQMTDEKYLWKDLTLDOAYGDAVENAKD 174  
 DB 123 RGPSSGSLHVGHLVFPIMTKWLQETFDVPLVQLTDEKTLMKDLKVEDAIKLGRENAD 182  
 QY 175 IIAAGFDINKTIFESDLDMYKMSGFGYKNNVKIQKHVTFNQVKGIFGFTSDICIGKISFP 234  
 DB 183 IYALGFDVKTFFIENNLFEVKGCPAMYQNIIRIQKCVTFNQVKGIFGFTSDICIGKISFP 242  
 QY 235 AIOAAPSFNSFPQIFRDRDTQCLIPCAIDDPYFRMTDVAIRIGYKPKALLHSTFEP 294  
 DB 243 AAQAAAPALISTPFPFIGNR-KVHCILPCALIDDPYFRMTDVAIRIGYKPKALLHSTFEP 301  
 QY 295 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRDTIEHRQPGNCDDVVSFM 354  
 DB 302 AIOGAQTMSASDPNSSIFLTDTAKQIKTKVNHAFSGGRDTIEHRQPGNCDDVVSFM 361  
 QY 355 YLTFPLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLAEHQARRKREVTDEIVKEFM 414  
 DB 362 LKTFPLEDDDKLEEVAVVASKGEMLTGEIKGLAVETLTPVEHQAAARKLITDEVLDKTF 421  
 QY 415 TPRKLSF 421  
 DB 422 ELRPLKF 428

RESULT 13  
 AAG23698  
 ID AAG23698 standard; Protein; 402 AA.  
 AC AAG23698;  
 XX  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 27101.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridization assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 01-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139763.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147202.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153707.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 15-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.



```

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-013624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

```

```

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157173.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 58.1%; Score 1304; DB 21; Length 426;  
 Best Local Similarity 60.3%; Pred. No. 5 4e-125; Indels 4; Gaps 2;  
 Matches 240; Conservative 67; Mismatches 87;

```

QY 29 DATEAR-EDFVDPMTVQTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRPHHFLR 86
DB 31 DEREASSEQVNVPMWEVSAKDGGKIDYDKLIDKFGCGRDDESLIDRVRQLTSRQPHVFLR 90
QY 87 RGIFFSHRDMNOVLDAVENKPPYLYTGRGSPSEAHVGHILPFIITKMLQDVFNPLVY 146
DB 91 RSVFFARDFNEILDYERGDKFFLYTGRGSPSEALHGLHPMTKYLOEAFKVLVY 150
QY 147 QMTDDEKYLKMDLTLDQAVDAVENAKDIIACGDFIDINKTIFISFDLDYMGSSGFYNNVYK 206
DB 151 QLTDDKSIYKMLSVESQRLARENKDIACGDFVTKTIFISFDYVYG--GAFYNNMK 208
QY 207 IQGHVTFNOVKGIIEFTSDCIGKISFPALIOAASFSSNFPQIFRDRDIOIQLIPALID 266
DB 209 VGCVTILNKAMGIFGSGEDPIAKISFPVQAVSPSSFPPIHPGKDNIRCLIPCALID 268
QY 267 DPFYFMTROVAPRIGYKPKALLHSTFFPALQAGQOTMSASDPSSIFLTDIAQOITKYN 326
DB 269 DPFYFMTROVAPRIGYKPKALLHSTFFPALQAGQOTMSASDPSSIFLTDIAQOITKYN 328
QY 327 KHAFFSGRPTIEHROFGGNCVDVGFMYLTFLEDDDKLEQIRKDYTGAMLTGELKTA 386
DB 329 RYAFSGGDSIEHRLGANLEVDIPVKYLSFLEDDSELEHKKEXGGRMLTGELVKR 388
QY 387 LIEVLOPLIAEHQARRKEVTDEIVKEFWTPRKLSFDFQ 424
DB 387 LIEVLOPLIAEHQARRKEVTDEIVKEFWTPRKLSFDFQ 424

```

Db 389 LIEVLTEIVEKRRARAATDEKMDAFMAVRPLPSKE 426

RESULT 15

ID AAB66931 standard; Protein; 424 AA.

AC AAB66931;

DT 12-APR-2001 (first entry)

DE Tryptophanyl-tRNA synthetase.

KM Tryptophanyl-tRNA synthetase; enzyme.

OS Candida albicans.

PN US6174713-B1.

PD 16-JAN-2001.

PF 16-JUN-1997; 97US-0876885.

PR 16-JUN-1997; 97US-0876885.

PA (CUBI-) CUBIST PHARM INC.

PI Shen X, Houman F;

XX WPI; 2001-201806/20.

DR N-PSDB; AAF55855.

PT New nucleic acids encoding Candida cytoplasmic tryptophanyl-tRNA

PT (cRNA) synthetases, useful for producing recombinant cRNA synthetases

PS and detecting inhibitor of Candida cRNA synthetase function

CC The present sequence is Candida albicans cytoplasmic tryptophanyl-tRNA

CC (cRNA) synthetase; cRNA synthetase is useful for producing recombinant

CC cRNA synthetases and detecting inhibitors of cRNA synthetase.

SQ Sequence 424 AA;

Query Match 54.3%; Score 1218.5; DB 22; Length 424;

Best Local Similarity 57.1%; Pred. No. 3.3e-116;

Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

31 TEAEDPVPMTVO---TSSAKIDYDKLIVFGSSKIDKELINRITERATGQRPHPFR 86

13 TESEBQKITPWEVGBAVNDCKSMGIDYDKLISQGTKHITETLERFKOYTGEEHPFLK 72

87 RGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKWLQDVFNPLVI 146

73 RGVFFSQRDRLDRLDYHGEPEFLYTGSGSSDSMHLGHMVPEFTKWLQEVFVPLVI 132

147 QMTDDDEKYLAK-DLITDQAVGDVAVENAKDIIACFPDINKTFTFSDLDMGMSGFTKAV 205

133 ELTDDDEKFLFKQDLTIDVKGFAENAKDIIAVGFNPENTFIFSDIQMG--GAFYENVV 190

206 KIQKHTVFNQVKGIFGFTSDSCIGKISFPALQAPFSNSFPQIFRDRDIOCLIPCAID 265

191 KTSRQITSTAKAVPFTSDSCIGKIHFAISIQATAPSSFPVGLPRTKPLIPCAID 250

266 QDPYFMTDVAAPRIYKPKALHSTFPALQAGTMSADPNSIPLDTAKQIKTKV 325

326 NKAFSGGSDTJEHROFGNCDVVSFWTLTFFLEDDDKLQIRKDYTSGLMTGELKK 385

311 NKAFSGGSDTJEHROFGNCDVVSFWTLTFFLEDDDKLQIRKDYTSGLMTGELKK 370

386 ALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSF 421

Db 371 ECITVLQEFVSAYQERRSKVDQVVEKFMKPKLVF 406

Search completed: July 10, 2003, 12:30:11  
Job time : 35.92 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:29:05 ; Search time 12.8485 Seconds  
(without alignments)  
970.956 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471  
Perfect score: 2246  
Sequence: 1 MSYKAAGBDYKADCPPGNP.....VTDEIVKEFMTPRKLSFDFQ 424

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	54.3	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	8.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	8.3	341	US-09-492-581-2	Sequence 2, Appl
4	185.5	8.3	341	US-09-425-666-2	Sequence 2, Appl
5	164.5	7.3	409	US-08-743-130A-39	Sequence 39, Appl
6	162.5	7.2	409	US-08-743-130A-2	Sequence 2, Appl
7	132.5	5.9	388	US-08-705-868-4	Sequence 4, Appl
8	132.5	5.9	388	US-09-123-615-4	Sequence 4, Appl
9	113.5	5.1	418	US-08-855-910-11	Sequence 11, Appl
10	109	4.9	377	US-09-352-990-28	Sequence 28, Appl
11	108	4.8	197	US-08-923-867-2	Sequence 2, Appl
12	108	4.8	197	US-08-928-100-4	Sequence 4, Appl
13	108	4.8	197	US-09-183-134-2	Sequence 2, Appl
14	108	4.8	197	US-09-492-581-4	Sequence 4, Appl
15	108	4.8	197	US-09-425-666-4	Sequence 4, Appl
16	104.5	4.7	370	US-08-415-593-45	Sequence 45, Appl
17	100.5	4.5	427	US-09-134-001C-5141	Sequence 5141, Ap
18	95	4.2	418	US-08-844-054-2	Sequence 2, Appl
19	95	4.2	418	US-09-347-333-2	Sequence 2, Appl
20	95	4.2	464	US-09-134-001C-4701	Sequence 4701, Ap
21	94.5	4.2	877	US-08-907-166-8	Sequence 8, Appl
22	93	4.1	607	US-09-134-001C-2994	Sequence 2994, Ap
23	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
24	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
25	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
26	90.5	4.0	920	US-08-101-593-4	Sequence 4, Appl
27	89.5	4.0	344	US-09-393-554-2	Sequence 2, Appl

28	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
29	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
30	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
31	89.5	4.0	920	US-08-101-593-2	Sequence 2, Appl
32	89.5	4.0	920	US-08-101-593-2	Sequence 2, Appl
33	89	4.0	719	US-08-082-849B-31	Sequence 31, Appl
34	88.5	3.9	501	PCT-US94-01624-31	Sequence 31, Appl
35	88	3.9	501	US-09-157-257-8	Sequence 8, Appl
36	88	3.9	606	US-08-883-514-3	Sequence 3, Appl
37	87.5	3.9	606	US-09-204-764-3	Sequence 3, Appl
38	87	3.9	855	US-08-890-865A-10	Sequence 10, Appl
39	87	3.9	428	US-08-331-625A-43	Sequence 43, Appl
40	87	3.9	402	US-09-494-151-43	Sequence 43, Appl
41	87	3.9	1452	US-09-134-001C-4511	Sequence 4511, Ap
42	87	3.9	1452	US-08-331-625A-2	Sequence 2, Appl
43	87	3.9	1452	US-09-494-151-2	Sequence 2, Appl
44	87	3.9	1452	PCT-US93-04384-18	Sequence 18, Appl
45	86.5	3.9	849	PCT-US93-04692-2	Sequence 2, Appl
				US-09-157-257-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-876-885-26  
Sequence 26, Application US/08876885  
Patent No. 6174713  
GENERAL INFORMATION:  
APPLICANT: Shen, Xiaoyu  
TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPHANYL-TRNA  
TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING  
TITLE OF INVENTION: SAME  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/876,885  
FILING DATE: 16-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPI97-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-876-885-26

Query Match 54.3%; Score 1218.5; DB 4; Length 424;  
Best Local Similarity 57.1%; Pred. No. 6.1e-127;  
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 31 TAAEDFDPFWQ---TSSAKGIDYDKLIVRGSSKIDKELNIRIRAGQRPHHFLR 86  
DB 13 TEEBOKITPMEVEGAVVDGKSMGIDYDKLISOFCKHTETETLRFQOVGEBEPFLK 72

QY 87 RGTFESHDMNVVLDAVENKKPEFLYTGRRGSSSEAMVGHILPEFTKMLQDVENVPLVI 146  
DB 73 RGFPPSSORDRLIDLYEHGEPFLYTGRRGSSSMHGHVPEFTKMLQDVENVPLVI 132  
QY 147 QMTDDEKYLWK-DLTLDQAVGDAVENAKDIACGFDINKTFIFSDLDYMGSSGFYKNV 205  
DB 133 ELTDDKELFLRHQGLTIDVGFPAENKDIIVGFNENFTIFSDLYMG--GAFYENVV 190  
QY 206 KIQKHTFNOVKIGTGFSDSDCTGKISFPALQAAPSNSPQJFRDRTIOCLIPCAID 265  
DB 191 RTSQITTSYAKAVGFTSDCTGKIHFAISQITAFSPSPDVLGDPKTPCLIPCAID 250  
QY 266 QDPFRMTBPAVPIGYPKPAHLSTFPALQAGTQKASDPSSIFLDTAKQIKTV 325  
DB 251 QDPFRMTBPAVADKTRTPKPAHLSTFPALQAGTQKASDPSSIFLDTAKQIKTV 310  
QY 326 NKAHFGSGRDTIEHRQFGNCVDVSPMYLTFPELDDDKLEQIRKDYTGAMLTGELK 385  
DB 311 NKAFSGGRATAEHRELQGNPEVDVAFQYLSFSPSYDEKLAQLEQYRKGEILSGEMK 370  
QY 386 ALIEVLOPLAEHQARKEVTDEIVKEMTPRKLSF 421  
DB 371 ECITVLOEFVSAVOERSKVDQVVERKMKHKLVF 406

## RESULT 2

US-08-928-100-2  
Sequence 2, Application US/08928100

Patent No. 6046174  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NO. 6046174e1 trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-2

Query Match 8.3%; Score 185.5; DB 3; Length 341;  
Best Local Similarity 24.0%; Pred. No. 5,46-12;

Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;  
QY 107 KPEYLYTGRRGSSSEAMVGHILPEFTKMLQDVENVPLVIQMTDDKTY-LMKDLTLDQAY 165  
DB 3 KPIIL-TGDRPTEK-LHIGHVYGSLSNR-----VLLQEBEKYMFVFLADQOAL 49  
QY 166 GDAVEN-----AKDIACGFDINKTFIF--SDLDYMGSSGFYKNV---KIQK 209  
DB 50 TDAKDPQITIVESTGVNADLYDLAVGLDPKNSITFIQSQIPBELAELSMYTNVLARLER 109  
QY 210 HTVFNQVKGIFGFTSDSDCTGKISFPALQAA--PSFNSPQJFRDRTIOCLIPCAID 267  
DB 110 NPTVKTEISQKGEBSIPGFLVYPIAQAADITAFKANY-----VPVGTQK 156  
QY 268 PYRMTRD-----VAPRIYKPAHLSTFPALQAGTQKASDPSSIF 313  
DB 157 PMIQREIYRSEFNANVNCVLEPFGIYPENE--RAGRLPGIDG--NARKSKS-LNNGIY 212  
QY 314 LTDTAKQIKTVKHAHFGSGRDTIEHRQFGNCVDVSPMYLTFE--LEDDDKLEQIRK 371  
DB 213 LADDAOTLRKQVMSMTDDPHIRVEDPGKIEG---MTFHYLDVGRPEDAQELDMKE 268  
QY 372 DYTSGAMLTGELKALIEVLOPLAEHQARKEVTDEI 409  
DB 269 RYQRGGLGVKTRKRYLLEILRELGPIRRRIEFADKM 306

## RESULT 3

US-09-492-581-2  
Sequence 2, Application US/09492581

Patent No. 6346409  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NO. 6346409e1 trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/492,581  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
APPLICATION NUMBER: 9619072.3  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-09-492-581-2

Query Match 8.3%; Score 185.5; DB 4; Length 341;  
Best Local Similarity 24.0%; Pred. No. 5,4e-12;  
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 107 KPEVLYTGSGPSEAMHGHILPIFTKMLQDVFNPLVIQMTDDEKY-LMKDLTLDDQAY 165  
DB 3 KPIILL-TGDRPTGK-LHIGHVYSLKNR-----VLLQEDKXDMFVFLADQQL 49  
QY 166 GDAVEN-----ANDIACGFDINKTFIF--SDLDYGMSSGFYKXNV---KIQK 209  
DB 50 THAKDPQIVTSGVNSVALDYLAAGLDPRKSTFIQSQIPELAEISMTYMLVSLARLER 109  
QY 210 HTFNVQVKIFGFTSDCTGKISFPALQAA--PSFSNSPQPIFRDRTDQCLIPCAIPDD 267  
DB 110 NPTVKTETISQKGFSGSIPFGFLVYPIAQADITAFKANY-----VPVGTQDK 156  
QY 268 PYFRMTRD-----VAPRIQYPKPALHSTFFPALQGAQTMSASDPNSSIF 313  
DB 157 PMIEQIREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG-NAMKSKS-LNNGIY 212  
QY 314 LDTDAKQIKTKVKNKAHFGSGGRDTEIEHROFGNCDVVSFMYLTFE--LEDDKLEQIRK 371  
DB 213 LADDDADTLRKXVMSWYTPDPIHVEDPQKIEGN---MVFHLYDVGREPDAQEIADMKX 268  
QY 372 DYTSGAMLGELKALIEVLOPIAHOARKEVTDEI 409  
DB 269 RYRGGLGADVTKRYLLETLERELGPIRERRIEFAKDM 306

## RESULT 4

US-09-425-666-2  
Sequence 2, Application US/09425666

GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6416976el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/425,666  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/928,100  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm, Edward R  
REGISTRATION NUMBER: 38,881  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-425-666-2

Query Match 8.3%; Score 185.5; DB 4; Length 341;  
Best Local Similarity 24.0%; Pred. No. 5,4e-12;  
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 107 KPEVLYTGSGPSEAMHGHILPIFTKMLQDVFNPLVIQMTDDEKY-LMKDLTLDDQAY 165  
DB 3 KPIILL-TGDRPTGK-LHIGHVYSLKNR-----VLLQEDKXDMFVFLADQQL 49  
QY 166 GDAVEN-----ANDIACGFDINKTFIF--SDLDYGMSSGFYKXNV---KIQK 209  
DB 50 THAKDPQIVTSGVNSVALDYLAAGLDPRKSTFIQSQIPELAEISMTYMLVSLARLER 109  
QY 210 HTFNVQVKIFGFTSDCTGKISFPALQAA--PSFSNSPQPIFRDRTDQCLIPCAIPDD 267  
DB 110 NPTVKTETISQKGFSGSIPFGFLVYPIAQADITAFKANY-----VPVGTQDK 156  
QY 268 PYFRMTRD-----VAPRIQYPKPALHSTFFPALQGAQTMSASDPNSSIF 313  
DB 157 PMIEQIREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG-NAMKSKS-LNNGIY 212  
QY 314 LDTDAKQIKTKVKNKAHFGSGGRDTEIEHROFGNCDVVSFMYLTFE--LEDDKLEQIRK 371  
DB 213 LADDDADTLRKXVMSWYTPDPIHVEDPQKIEGN---MVFHLYDVGREPDAQEIADMKX 268  
QY 372 DYTSGAMLGELKALIEVLOPIAHOARKEVTDEI 409  
DB 269 RYRGGLGADVTKRYLLETLERELGPIRERRIEFAKDM 306

## RESULT 5

US-08-743-130A-39  
Sequence 39, Application US/08743130A

GENERAL INFORMATION:  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Tao, Nianjun  
APPLICANT: Tao, Jianshi  
TITLE OF INVENTION: CANDIDA TYROSYL-ERNA SYNTHETASE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,130A  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPI95-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids

TYPE: amino acid  
STRANDEDNESS: linear  
US-08-743-130A-39

Query Match  
Best local Similarity 20.7%; Pred. No. 1.6e-09;  
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

7.3%; Score 164.5; DB 2; Length 409;

99 VLDAYENK-KPFLYTGRRGSSSEAMVGHILPFI-----FTKMLQDV-----F 140  
27 IKVLEKENRPVXIYWGTAFTGKP-HCGYFVPMIKAHFLKAGCEVTLADLHAFLDNM 85  
141 NVPLVIGMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFINDKTIFFSDLDYMGSSGF 200  
86 KAPLEVVKYRAKYEFVVKAILKSNIVPIERLKFVVGSSYQKGDYV---MDLFLKSNIV 142  
201 YKAVVK-----IQKVTFNQVKGIFGFTSDICIGISFPALQAPAFSNSFPQIFRDRTD 255  
143 SQMDARAGADVVKQVANPLSLGI-----YPLMQA-----IDEEHLG 180  
256 IQCLIPCAIDQDPYFMTDVAIRIGYKPKALHSTFFPALQAGQTMASADPNISIFLT 315  
181 VDAQFG-GVDQKRFVLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDII 238  
316 DTKAKQITKTKNKAIFSGG--RDT-----IEEHROGCGN 346  
239 EEPKVVKKKVNAYCAPGELKDNGLAFIEVYIQPIAELKTGVEGAFKLDIDREKYG- 297  
347 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALI-----EVLQPLAIEHQARR 402  
298 ---PLSY-----DSIEQLKADFVQDKLAPPLDKSGVADKINLAPIRAEFESS- 343

QY 403 KEVTDEIVKEFMTPRK 418  
DB 344 -----EEFOVAQK 351

RESULT 6  
US-08-743-130A-2  
Sequence 2, Application US/08743130A  
Patent No. 5871987  
GENERAL INFORMATION:  
APPLICANT: Sasanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Tao, Nianjun  
APPLICANT: Houman, Fariba  
TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743.130A  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: C9195-12  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
US-08-743-130A-2

Query Match  
Best local Similarity 20.7%; Pred. No. 2.7e-09;  
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

7.2%; Score 162.5; DB 2; Length 409;

99 VLDAYENK-KPFLYTGRRGSSSEAMVGHILPFI-----FTKMLQDV-----F 140  
27 IKVLEKENRPVXIYWGTAFTGKP-HCGYFVPMIKAHFLKAGCEVTLADLHAFLDNM 85  
141 NVPLVIGMTDDEKYLMDLTLDOAYGDVAVENAKDIIACGFINDKTIFFSDLDYMGSSGF 200  
86 KAPLEVVKYRAKYEFVVKAILKSNIVPIERLKFVVGSSYQKGDYV---MDLFLKSNIV 142  
201 YKAVVK-----IQKVTFNQVKGIFGFTSDICIGISFPALQAPAFSNSFPQIFRDRTD 255  
143 SQMDARAGADVVKQVANPLSLGI-----YPLMQA-----IDEEHLG 180  
256 IQCLIPCAIDQDPYFMTDVAIRIGYKPKALHSTFFPALQAGQTMASADPNISIFLT 315  
181 VDAQFG-GVDQKRFVLAENLPSIGYKRAHLMNPMVPGI-GQGGKMSASDPNSKIDII 238  
316 DTKAKQITKTKNKAIFSGG--RDT-----IEEHROGCGN 346  
239 EEPKVVKKKVNAYCAPGELKDNGLAFIEVYIQPIAELKTGVEGAFKLDIDREKYG- 297  
347 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALI-----EVLQPLAIEHQARR 402  
298 ---PLSY-----DSIEQLKADFVQDKLAPPLDKSGVADKINLAPIRAEFESS- 343

QY 403 KEVTDEIVKEFMTPRK 418  
DB 344 -----EEFOVAQK 351

RESULT 7  
US-08-705-868-4  
Sequence 4, Application US/08705868  
Patent No. 5885798  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,868  
FILING DATE: Filed Herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749

```

REFERENCE/DOCKET NUMBER: PF-0117 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1184699
US-08-705-868-4

Query Match
Best Local Similarity 20.8%; Pred. No. 5.3e-06;
Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

114 GRGPS-SEAMVGHILPIFTKMLQDVNVPLVIOQMTDE--KYLMDLTLDO--AYGD 167
2 GDAPSPEEKH-----LITRNLOEVLGEBEKLEKILKERELKIYWGATTCGPHVAVFV 54
168 AVENAKDITACGPDINKTFIFSDL-----DYMGMSSGFYKVVVKIQKH--VTF 213
55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELRLRSYENYIKAMLESIGVPL 112
214 NOVKGIFG-----FTDSD-----CIGKISFPAIOA 239
113 EKLKFKGTGYOLSKETLDVYRLSSVYQHDSKKAQAEVVKQVEHPHLSGLIYPLQO- 171
240 PFSNPSFPIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPRKALHSTFFPALOGA 299
172 --LDEBYLKV-----DAQF--GGIDQRKIFTPAEKYLPALGYSKRVHLMNPMVPGLTG- 220
300 QTKMSASDPNSSIFLTDPAKQIKTVNKAFA-----SGGRDTIEH----- 340
221 -SKMSSSEESKIDLDKREKDYKKLKK-AFCCEPGVNNNGVLSFKHVLPLKSEFVYL 278
341 --ROFGNCDVDSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEH 398
279 RDEKMGNN-----KTYTAYVD-----LEKDPAAEVVHPGDLKNSVEVALNKLL----- 321
399 QARRKEVTDIEYKEPMTF--RKLS 420
322 -----DPIREKNTPALKKLA 337

RESULT 8
US-09-123-615-4
Sequence 4, Application US/09123615
Patent No. 6090377
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Murty, Lynn B.
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/09/123,615
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,868
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0117 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1184699
US-09-123-615-4

Query Match
Best Local Similarity 20.8%; Pred. No. 5.3e-06;
Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

114 GRGPS-SEAMVGHILPIFTKMLQDVNVPLVIOQMTDE--KYLMDLTLDO--AYGD 167
2 GDAPSPEEKH-----LITRNLOEVLGEBEKLEKILKERELKIYWGATTCGPHVAVFV 54
168 AVENAKDITACGPDINKTFIFSDL-----DYMGMSSGFYKVVVKIQKH--VTF 213
55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELRLRSYENYIKAMLESIGVPL 112
214 NOVKGIFG-----FTDSD-----CIGKISFPAIOA 239
113 EKLKFKGTGYOLSKETLDVYRLSSVYQHDSKKAQAEVVKQVEHPHLSGLIYPLQO- 171
240 PFSNPSFPIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPRKALHSTFFPALOGA 299
172 --LDEBYLKV-----DAQF--GGIDQRKIFTPAEKYLPALGYSKRVHLMNPMVPGLTG- 220
300 QTKMSASDPNSSIFLTDPAKQIKTVNKAFA-----SGGRDTIEH----- 340
221 -SKMSSSEESKIDLDKREKDYKKLKK-AFCCEPGVNNNGVLSFKHVLPLKSEFVYL 278
341 --ROFGNCDVDSFMYLTFPLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEH 398
279 RDEKMGNN-----KTYTAYVD-----LEKDPAAEVVHPGDLKNSVEVALNKLL----- 321
399 QARRKEVTDIEYKEPMTF--RKLS 420
322 -----DPIREKNTPALKKLA 337

RESULT 9
US-08-855-910-11
Sequence 11, Application US/08855910
Patent No. 6221640
GENERAL INFORMATION:
APPLICANT: Tao, Jianshi
APPLICANT: Sassanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Avrukh, Anthony S.
APPLICANT: Yu, Russell V.
APPLICANT: Nair, Shantia
TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/855,910  
 FILING DATE: 14-MAY-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook, David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-08  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 861-6240  
 TELEFAX: (781) 861-9540  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 418 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-855-910-11

Query Match 5.1%; Score 113.5; DB 4; Length 418;  
 Best Local Similarity 20.5%; Pred. No. 0.00079;  
 Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;  
 111 LTYGRGSSAMAVGHILPFIETKMLQDVFNPLVLT----- 146  
 33 LYGVDPDGTSMHIGLIPFMMKRFQLAGHHPYLLGGGTGTIGDPSGRTTREVLTQME 92  
 147 -----OWTDEKYLW-KDLTDQAVGAVEN-----AKQIIA 177  
 93 AVQHNVDLSNOMKKLFGKNAEVTMANNVYDMLSELSDLPLRDYGNFNNVMTLAKDIIA 152  
 178 CGEDINKTIFSDLDYVMSS-GEYKNNVKIQKVTENVKG----- 218  
 153 S--RLSGISFTERTYQILQSIDFY--TLHKKNIQQLGADQMGNTIAGDLIRKKE 207  
 219 -----IGFT-----DSDICIKISFPALQAPSFNSFP--QIFPRDRDIOCLIFCAI 264  
 208 GPEAKYFGLTIPMLKADGTRKFGTAGAIWLDPKKTSPEFYQFWMLOD----- 258  
 265 DQDPYFMTDVAFRIGYPRPALHSTFPALQAGOTKMSASDPNSSIFLDTAKQIKTK 324  
 259 -----RDV--IKYK-----FTFLDKER-----IDALAEKYEKE 286  
 325 VNKAFFSGGDTIEHROFGGNCVDVSMYLLFFLEDDKLEQIRKDYSGAMLTGELK 384  
 287 PGRK--EAQRLEAE-----VTRFVHDDALEBAQK--ISEALFSGNIK 326  
 385 KALIEVQ-----PLAEHQARKEVTDEIVKEFMTPRK 418  
 327 DLTIETIEQLEHVFV-EITKDAKNIVDWLVDTIEPSK.365

RESULT 10  
 US-09-352-990-28  
 Sequence 28, Application US/09352990  
 Patent No. 6255090  
 GENERAL INFORMATION:  
 APPLICANT: Farnodu, Layo O.  
 APPLICANT: Orozco, Buddy  
 APPLICANT: Rafalaki, Antoni  
 TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
 FILE REFERENCE: BB-1191

CURRENT APPLICATION NUMBER: US/09/352,990  
 CURRENT FILING DATE: 1999-07-14  
 EARLIER APPLICATION NUMBER: 60/092,866  
 EARLIER FILING DATE: July 15, 1998  
 NUMBER OF SEQ ID NOS: 29  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 28  
 LENGTH: 377  
 TYPE: PRT  
 ORGANISM: Synechocystis sp.  
 US-09-352-990-28

Query Match 4.9%; Score 109; DB 4; Length 377;  
 Best Local Similarity 21.0%; Pred. No. 0.0021;  
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;  
 91 FSHRDMNQVLDAVENKKPFLVYTGKPSSEAMHVGHLIFITKMLQDVFNPLVYQMTD 150  
 33 FAHRSTTAM-----DKPRIL-SGVQPTGN-LHLGNVLAIR-SWVEQ-----QOHY 75  
 151 DEKYMKDL-----TLQAVGDAVENAKDIIAGCFDINKTPIF----- 188  
 76 DNFECVVDLHAITVPHNPOTLQO--DTTIALYIACIIDIQSTIFVQSHVAHSELA 132  
 189 -----SDLDYMGSSGFYKNNVKIQKVTENVQKGFQPTSDCIGKISFPALQAPSF 242  
 133 WLNCVTPLNMLERMQFKEKAVKQGENVS-----VGLLDYFVLMAA----- 174  
 243 SNSFPQIFRDRDIOCL-----IPCAIDDPYFRMTD-----ARIGYPKRA 286  
 175 -----DILLYDAKVPVGEDOKQHELTRDIYIRINDKRGREDAVLLKPEPL 222  
 287 L-LHSTFPALQAGOTKMSASDPN--SIFLDTAKQIKTKYKNAFSGGRDTIEHROF 343  
 223 IRKEGARVNSLADGTYKMSKSDSELSRLNLDPRPMIKKVKK----- 266  
 344 GNCNDVDSFMYLTFPLEDDDKLEQIRKDYSGAMLTGELKAL----- 387  
 267 ---CKTDPO-RGLMF---DDPERBECHNLLTYLLTSLNQTKEAVQBCAEMWGQFKPL 319  
 388 -----IEVLOPLAEHQARKEVTDEIVK 411  
 320 TETAIALPPIQAKYAEIILADRGL-DRILIQ 349

RESULT 11  
 US-08-923-867-2  
 Sequence 2, Application US/08923867  
 Patent No. 5851809  
 GENERAL INFORMATION:  
 APPLICANT: Lawlor, Elizabeth  
 TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SmithKline Beecham Corporation  
 STREET: 709 Swedeland Road  
 CITY: King of Prussia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19406-0939  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/923,867  
 FILING DATE: 04-SEP-1997  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 9619072.3  
 FILING DATE: 12-SEP-1996  
 ATTORNEY/AGENT INFORMATION:

NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-867-2

Query Match  
Best Local Similarity 25.3%; Pred. No. 0.00092;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRTRD-----VAPRIGYPRKALLHSTFPALQGAQTKMSA 305  
DB 5 VEVGTDDQKPMIEQTRIVSFNNAYNCVILVEPEGIYPENE--RAGRLEGLDG-NAKMSK 61

QY 306 SDPNSSIFLPTDPAKQIKTKVKNKHAFGSGRDTIEHRQFGNCDVDSFWYLTFF--LEDD 363  
DB 62 S-LNNGIYIADADADTLRKVKVMSYTDPDHIREVDPKIEGN---WVFHYLDVFGRPEDA 116

QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDEI 409  
DB 117 QRIADMKERYQRGGLGDVYTKRYLLEIRELGPRIERRIEPAKDM 162

RESULT 12  
US-08-928-100-4  
Sequence 4, Application US/08928100  
Patent No. 6046174  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NO. 6046174el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-4

Query Match  
Best Local Similarity 4.8%; Score 108; DB 3; Length 197;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRTRD-----VAPRIGYPRKALLHSTFPALQGAQTKMSA 305  
DB 5 VEVGTDDQKPMIEQTRIVSFNNAYNCVILVEPEGIYPENE--RAGRLEGLDG-NAKMSK 61

QY 306 SDPNSSIFLPTDPAKQIKTKVKNKHAFGSGRDTIEHRQFGNCDVDSFWYLTFF--LEDD 363  
DB 62 S-LNNGIYIADADADTLRKVKVMSYTDPDHIREVDPKIEGN---WVFHYLDVFGRPEDA 116

QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLQPLIAEHQARKEVTDEI 409  
DB 117 QRIADMKERYQRGGLGDVYTKRYLLEIRELGPRIERRIEPAKDM 162

RESULT 13  
US-09-183-134-2  
Sequence 2, Application US/09183134  
Patent No. 6165759  
GENERAL INFORMATION:  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,134  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/923,867  
FILING DATE: 04-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-183-134-2

Query Match  
Best Local Similarity 4.8%; Score 108; DB 4; Length 197;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```
QY      260 PCALIDODPYFMETD-----VPRICYPKRALHSHFFPALOAGRKMSA 305
        :|::|||:
Db       5 VPGVDQKMEIQTREIVASFNNAANCVLVEPEGIIPENE--RAERLPGLDS-NAKMSK 61
        :|::|||:
QY      306 SDPNSSIFLTPDAKOIKTYKNKAHAFSGGRDTIEHRQFGNCVDVSPFWLTFF--LEDD 363
        :|::|||:
Db       62 S-LNNGIYIADADPTLRKKVMGMYTDPDHIREEDPGKITEN---WVFHYLDVFGRPEDA 116
        :|::|||:
QY      364 DKLEQIRKDQYTGAMLTGELKTALLIEVLPLAEHQARRKEVTDEI 409
        :|::|||:
Db      117 QEADMKERIYRGGLGDVTKRYLIEILERELOCFIERRIEFKMD 162
```

RESULT 14  
US-09-492-581-4  
: Sequence 4, Application US/09492581

Patent No. 6346409  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danlle  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6346409el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/452,561  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm, Edward R  
REGISTRATION NUMBER: 38, 891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-492-581-4

Query Match	4.8%	Score 108	DB 4	Length 197
Best Local Similarity	25.3%	Pred. No. 0.00092		
Matches 42	Conservative 27	Mismatches 73	Indels 24	Gaps 6

Ddb Qy 260 IPGAIIDDDPRFRMTD-----VAPRGYPKALHLSSTFFPLQGAQTMSA 305  
:::|::|  
5 VPGVTGDKPMIEQTRFIVRSFNNAAYNCDLVEPEGIYPPNE--RGRLLPGLDG-NAKMSK 61

QY 306 SDPNSIIFLDTAKQKTKVKAKHAFSGGRDTIEHHRGFGGNCDDVAFMYLTFE--LEDD 363

Db 62 S-LNNGIYLLDDADDTLRKKKSMYTPDPIHVEDPGKIEGN---MVFHYLDFVGRBEDA 116

```

Oy      364 DKLEQIRKDYTSGLMTGELKKALIEVLQPLIAHQARKKEVTDEI 409
        :: :: | | : : | : | : : | |
Db      117 QEIADMKERYQRGGLGDVTKRYLLEILERELGPIRRIEFAKDM 1622

```

RESULT 15  
US-09-425-666-4  
; Sequence 4, Application US/09425666

Patent No. 6416976  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6416976el ttps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/425.666  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/928,100  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-425-666-4

Query Match	4.8%;	Score 108;	DB 4;	Length 197;
Best Local Similarity	25.3%;	Pred. No. 0.00092;		
Matches 42;	Conservative 27;	Mismatches 73;	Indels	

```

QY      260  IPCALDQDDPYFRMTRD-----VAPRIGYKPKPALHSTFEPA
           :| | | | | :| | | | | :| | | | |
Db      5  VPVGTDQKPMIEQTREIVRSFNNAYNCDLVEPEGIYPENE--RAGRLPG

```

```

QY      306 SDPNSIFLTDTAKQIKTKYNKHAFSGGRDTIEHRQFGNCNDVSVSMY
          | | : | : | : | : | : | : | : | : | : | : | : |
Db      62 S-LNNGIYLADADATLRKKYMSMYTDPDHRVEDPGKIEGN---MVFHY

```

```
QY      364 DKLEGIKKDIYSGAMLTGGKALLIEVLQPFLAEHQARRKEVIDEL    409  
       ::::| | : :|||: : ||  
Db     117 QEIDMKERYÖRGGLGVTKRYLLEILERELGPIRERRIEFAKDM   162
```

Search completed: July 10, 2003, 12:32:50  
Job time : 15.8485 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:28:00 ; Search time 16.4461 Seconds  
(without alignments)  
2478.466 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEMTPRKSPDFQ 424

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73:\*

1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1 A41706	tryptophan-tRNA 11
2	2134.5	95.0	475	1 YWBO	tryptophan-tRNA 11
3	2041	90.9	475	1 YMRBPR	tryptophan-tRNA 11
4	2032	90.5	481	2 S50053	tryptophan-tRNA 11
5	1210	53.9	395	2 S58157	hypothetical prote
6	1163	51.8	432	2 S51901	tryptophan-tRNA 11
7	907	40.4	386	2 C90190	tryptophan-tRNA
8	803	35.8	385	2 C75020	tryptophan-tRNA
9	626.5	27.9	301	2 G71206	tryptophan-tRNA 11
10	534.5	23.8	380	2 G84373	tryptophan-tRNA 11
11	409.5	18.2	370	2 F64476	tryptophan-tRNA 11
12	397.5	17.7	364	2 E69131	tryptophan-tRNA 11
13	386	17.2	134	2 T43806	tryptophan-tRNA 11
14	370.5	16.5	420	2 E69461	tryptophan-tRNA 11
15	353	15.7	374	2 D72477	probable tryptoph
16	274.5	12.2	513	2 F84371	tryptophan-tRNA
17	192	8.5	323	2 H69346	tyrosyl-tRNA synth
18	188	8.4	364	2 E72512	probable tyrosyl-t
19	185.5	8.3	341	2 D95260	tryptophan-tRNA
20	185.5	8.3	341	2 G98125	tryptophan-tRNA 11
21	174.5	7.8	341	2 B86633	tryptophan-tRNA 11
22	172	7.7	394	2 A85999	tyrosine-tRNA 11g
23	170.5	7.6	366	2 S75410	tyrosine-tRNA 11g
24	169	7.5	395	2 H70385	tryptophan-tRNA 11
25	167	7.4	351	2 E75438	tryptophan-tRNA
26	166.5	7.4	346	2 B71496	tryptophan-tRNA 11
27	164.5	7.3	895	2 A86410	protein F3M18.22
28	163.5	7.3	353	2 E70100	tryptophan-tRNA 11
29	163	7.3	460	2 C84750	probable tyrosyl-t

30	159.5	7.1	337	2 F71300	tryptophan-tRNA 11
31	158.5	7.1	375	2 B75072	tyrosyl-tRNA synth
32	158	7.0	346	2 C91654	tryptophan-tRNA
33	154.5	6.9	102	2 T44994	probable tryptoph
34	152	6.8	344	2 H86590	tryptophan-tRNA
35	152	6.8	344	2 C72034	tryptophan-tRNA 11
36	149.5	6.7	327	2 C84374	tyrosyl-tRNA synth
37	149.5	6.7	408	2 T03741	probable tyrosine-
38	146	6.5	339	2 E64676	tryptophan-tRNA 11
39	146	6.5	375	2 F71093	tyrosine-tRNA 11g
40	145	6.5	337	2 A11066	tryptophan-tRNA 11
41	138	6.1	319	2 H69102	tyrosine-tRNA 11g
42	135.5	6.0	328	2 C72370	tryptophan-tRNA 11
43	135	6.0	365	2 E82052	tryptophan-tRNA
44	132.5	5.9	343	2 S73024	tryptophan-tRNA 11
45	132	5.9	328	1 YWBSF	tryptophan-tRNA 11

#### ALIGNMENTS

##### RESULT 1

A41706  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homo)  
C:Species: Homo sapiens (man)  
C>Date: 19-May-2000 #sequence revision 19-May-2000 #text change 03-Jun-2002  
C:Accession: A41633; A41706; S19246; J06076; J0533; S26287  
R:Elekter, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma  
A:Reference number: A41633; MUID:92107982; PMID:1763065  
A:Accession: A41633  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophan-tRNA synthetase expression in human fibroblasts.  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A417

A;Residues: 1-212, 'GD', 215-471 <PRO2>  
 A;Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368  
 A;Experimental source: /fibroblast  
 C;Genetics:  
 A;Gene: GDB:MARS; IFF53  
 A;Cross-references: GDB:119632; OMIM:191050  
 A;Map position: 14q23-14q31  
 A;Intons: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C;Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C;Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F;19-64/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 1.8e-173; Indels 0; Gaps 0;  
 Matches 422; Conservative 0; Mismatches 2;

QY 1 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAATEAEEDFVDPWVQTSSAKGIDYDKLIVRF 60  
 DB 48 MSYKAAAGEDYKADCPGPNPAPTSNHPDPAATEAEEDFVDPWVQTSSAKGIDYDKLIVRF 107  
 QY 61 GSKIDELINIRERATGQRPHRLRGIFFSHRDNNQVLDAYNKKPFYLYTGRGSSSE 120  
 DB 108 GSKIDELINIRERATGQRPHRLRGIFFSHRDNNQVLDAYNKKPFYLYTGRGSSSE 167  
 QY 121 AMVGHILPFIPTKMLQDVFNPLVIQMTDEKYLMDLTLDOAYGAVENAKDIACGFD 180  
 DB 168 AMVGHILPFIPTKMLQDVFNPLVIQMTDEKYLMDLTLDOAYGAVENAKDIACGFD 227  
 QY 181 DINKTFISDLIDYMGSSGFYKVVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAPS 240  
 DB 228 DINKTFISDLIDYMGSSGFYKVVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAPS 287  
 QY 241 SFNSFPQIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPFALLHSTFFPALQAGT 300  
 DB 288 SFNSFPQIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPFALLHSTFFPALQAGT 347  
 QY 301 TKMSASDPNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFPLE 360  
 DB 348 TKMSASDPNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFPLE 407  
 QY 361 EDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARRKEVTDEIYKEFMTPKLS 420  
 DB 408 EDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARRKEVTDEIYKEFMTPKLS 467  
 QY 421 FDFQ 424  
 DB 468 FDFQ 471

## RESULT 2

YMRBO  
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - bovine

N;Alternate names: tryptophanyl-tRNA synthetase  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text\_change 03-Jun-2002  
 C;Accession: A40279; UN0354; S10450; S14540  
 R;Garret, M.; Pajot, B.; Trezeuet, V.; Labouesse, J.; Metle, M.; Gandar, J.C.; Bendet  
 Biochemistry 30, 7809-7817, 1991  
 A;Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic  
 A;Reference number: A40279; MUID:9132348; PMID:1907847  
 A;Accession: A40279  
 A;Molecule type: mRNA  
 A;Residues: 1-475 <GAR>  
 A;Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA3079.1; PID:g163799  
 A;Experimental source: pancreas  
 A;Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue  
 R;Zargartava, T.A.; Kovalova, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.  
 Biorg. Khim. 15, 1307-1311, 1989  
 A;Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca  
 A;Reference number: UN0354; MUID:90211408; PMID:2631664  
 A;Accession: UN0354  
 A;Molecule type: protein

A;Residues: 112-124;282-287, 'N', 288, 'F', 289-292, 'Q', 293-294, 'IR', 336-353; 423-441, 443-449,  
 A;Experimental source: liver  
 A;Note: this paper is in Russian  
 C;Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C;Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F;24-69/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;  
 Best Local Similarity 95.0%; Pred. No. 1.3e-165; Indels 1; Gaps 1;  
 Matches 402; Conservative 9; Mismatches 11;

QY 2 SYKAAAGEDYKADCPGPNPAPTSNHPDPAATEAEEDFVDPWVQTSSAKGIDYDKLIVRF 61  
 DB 54 SYKAAAGEDYKADCPGPNPAPTSNHPDPAATEAEEDFVDPWVQTSSAKGIDYDKLIVRF 113  
 QY 62 SSKIDELINIRERATGQRPHRLRGIFFSHRDNNQVLDAYNKKPFYLYTGRGSSSE 121  
 DB 114 SSKIDELINIRERATGQRPHRLRGIFFSHRDNNQVLDAYNKKPFYLYTGRGSSSE 173  
 QY 122 MHVGHILPFIPTKMLQDVFNPLVIQMTDEKYLMDLTLDOAYGAVENAKDIACGFD 181  
 DB 174 MHVGHILPFIPTKMLQDVFNPLVIQMTDEKYLMDLTLDOAYGAVENAKDIACGFD 232  
 QY 182 INKTFISDLIDYMGSSGFYKVVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAPS 241  
 DB 233 INKTFISDLIDYMGSSGFYKVVVKIQKVTENQVKGIFGFTSDCIGKISFPAIOAPS 292  
 QY 242 SFNSFPQIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPFALLHSTFFPALQAGT 301  
 DB 293 SFNSFPQIFRDRDIOCLIPCAIDODPYFRMTDVAPRIGYPFALLHSTFFPALQAGT 352  
 QY 302 KMSASDPNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFPLE 361  
 DB 353 KMSASDPNSIFLITDPAKQIKTKVNGHAFSGGSDTIEHRQFGNCDDVVSFMYLTFPLE 412  
 QY 362 DDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARRKEVTDEIYKEFMTPKLS 421  
 DB 413 DDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAEHQARRKEVTDEIYKEFMTPKLS 472  
 QY 422 DDFQ 424  
 DB 473 DDFQ 475

## RESULT 3

YMRBRP

tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit

N;Alternate names: tryptophanyl-tRNA synthetase  
 C;Species: Oryctolagus cuniculus (domestic rabbit)  
 C;Date: 30-Sep-1992 #sequence revision 13-Feb-1998 #text\_change 03-Jun-2002  
 C;Accession: A35904; S37396  
 R;Lee, C.C.; Craig, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990  
 A;Title: Cloning and expression of a mammalian peptide chain release factor with sequence  
 A;Reference number: A35904; MUID:90239043; PMID:2185472  
 A;Accession: A35904  
 A;Molecule type: mRNA  
 A;Residues: 1-475 <LBR>  
 A;Cross-references: GB:M33460  
 R;Folova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Drugeon, G.; McCaughan, K.K.;  
 EMBL J. 12, 4013-4019, 1993  
 A;Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase are  
 A;Reference number: S37396; MUID:94009008; PMID:8404667  
 A;Accession: S37396  
 A;Molecule type: mRNA  
 A;Residues: 166-177 <PRO>  
 C;Genetics:  
 A;Gene: WRS  
 C;Complex: homodimer [validated, MUID:94009008]  
 C;Function:  
 A;Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent formati  
 A;Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain release  
 C;Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protein h  
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>  
 F:174-177/Region: ATP-binding motif (HKGH)

Query Match 90.9%; Score 2041; DB 1; Length 475;  
 Best Local Similarity 89.8%; Pred. No. 5,1e-158;  
 Matches 380; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

2 SYKAAAGBYKADCPGPNAPTSNHGPDATBAEDPVDPWTQTSSAKGIDYDKLIYRF 61  
 |||  
 53 SYKAAAGBYKADCPGPNAPTSNHGPDATBAEDPVDPWTQTSSAKGIDYDKLIYRF 112  
 |||  
 62 SKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 121  
 |||  
 113 SKKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 172  
 |||  
 122 MHVGHILPFIPTKMLQDVFNPLVIOQTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 181  
 |||  
 173 MHVGHILPFIPTKMLQDVFNPLVIOQTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 232  
 |||  
 182 INKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDICIGKISFPALQAA 241  
 |||  
 223 INKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDICIGKISFPALQAA 292  
 |||  
 242 FSNSEFQIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFFPALQAA 301  
 |||  
 293 FSNSEFQIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFFPALQAA 352  
 |||  
 302 KMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 361  
 |||  
 353 KMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 412  
 |||  
 362 DDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARKKEVTDEIYKEMTPRKLS 421  
 |||  
 413 DDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARKKEVTDEIYKEMTPRKLS 472  
 |||  
 422 DFG 424  
 |||  
 473 HQ 475

RESULT 4  
 S50053  
 C:Species: Mus musculus (house mouse)  
 C:Date: 07-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 03-Jun-2002  
 C:Accession: S50053; S50052; I49391; S31461; S31462  
 J:Paot, B.; Sarger, C.; Bonnet, J.; Garret, M.  
 J: Mol. Biol. 242, 599-603, 1994  
 A:Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthe  
 A:Reference number: S50052; M01D:95018226; PMID:7932716  
 A:Accession: S50053  
 A:Molecule type: mRNA  
 A:Residues: 1-481 <PAJ>  
 A:Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438  
 A:Genetics: LSF  
 A:Note: Intron position was determined by sequencing of genomic DNA  
 A:Accession: S50052  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <PAM>  
 A:Cross-references: EMBL:X69656; NID:g55435; PIDN:CAA49347.1; PID:g55436  
 A:Genetics: SSF  
 R:Kiselev, L.L.  
 Biochimie 75, 1027-1039, 1993  
 A:Title: Mammalian tryptophanyl-tRNA synthetases.  
 A:Reference number: I49391; M01D:94257729; PMID:7515282  
 A:Accession: I49391  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-481 <RES>  
 A:Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438  
 C:Genetics: <LSF>  
 A:Gene: WRS

A:introns: 475/2  
 A:Note: the list of introns may be incomplete; clone W13  
 C:Genetics: <SSF>  
 A:Gene: WRS

A:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C:Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis  
 F:1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted  
 F:1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted  
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 90.5%; Score 2032; DB 2; Length 481;  
 Best Local Similarity 89.6%; Pred. No. 2.8e-157;  
 Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

1 MSYKAAAGBYKADCPGPNAPTSNHGPDATBAEDPVDPWTQTSSAKGIDYDKLIYRF 60  
 |||  
 52 MSYKAAAGBYKADCPGPNAPTSNHGPDATBAEDPVDPWTQTSSAKGIDYDKLIYRF 111  
 |||  
 61 GSKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 120  
 |||  
 112 GSKIDKELINIRERATGQRPHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSPA 171  
 |||  
 121 AMHGHILPFIPTKMLQDVFNPLVIOQTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 180  
 |||  
 172 AMHGHILPFIPTKMLQDVFNPLVIOQTDDEKYLMDLTLDQAYGDAVENAKDIIACGF 231  
 |||  
 181 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDICIGKISFPALQAA 240  
 |||  
 222 DINKTFIFSDLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDICIGKISFPALQAA 291  
 |||  
 241 SFSNPFQIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFFPALQAA 300  
 |||  
 292 SFSNPFQIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYKPKALHSTFFPALQAA 351  
 |||  
 301 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 360  
 |||  
 352 TKMSASDPNSSIFLTDPAKQIKTKVKNKAFSGGRDTIEHRQFGNCVDVDSFMYLTFFLE 411  
 |||  
 361 EDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARKKEVTDEIYKEMTPRKLS 420  
 |||  
 412 EDDKLEQIRKDYTSGLMTELKALIEVLQPLIAEHQARKKEVTDEIYKEMTPRKLS 471  
 |||  
 421 DFG 424  
 |||  
 472 HQ 475

RESULT 5  
 S58157  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 10-Dec-1999  
 C:Accession: S58157; T38561  
 R:Gentles, S.; Churcher, C.M.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: S58145  
 A:Accession: S58157  
 A:Molecule type: DNA  
 A:Residues: 1-395 <GEN>  
 A:Cross-references: EMBL:Z50142; NID:g1052783; PIDN:CA90500.1; PID:g1052796  
 R:Gentles, S.; Churcher, C.M.; Barrett, B.G.; Rajandream, M.A.; Wain, S.V.  
 submitted to the EMBL Data Library, July 1995  
 A:Accession: T38561  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-395 <GB2>  
 A:Cross-references: EMBL:Z50142; PIDN:CA90500.1; GSPDB:GN00066; SPDB:SPAC2F7.13C  
 A:Experimental source: strain 972h-; cosmid c2f7  
 C:Genetics: SPDB:SPAC2F7.13C  
 A:Gene: SPDB:SPAC2F7.13C

A:Map position: 1

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 53.9%; Score 1210; DB 2; Length 395;

Best Local Similarity 59.6%; Pred. No. 1.5e-90;

Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 34 EEDFVDPWVQTS-----SAGIDYDKLIVRGSSKIDKELINRERATGGRPHFARGI 89  
 DB 4 EEOIVTMDVKGSLVDEBEGKIDYERLIVQGTIKRITPEOLERPEKLTGKKPHLLRGA 63  
 QY 90 FFSHRDNNQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFITKMLQDVFNPLVIQMT 149  
 DB 64 FFSHRDPMILDRERQKKPFYLYTGRGSSSMHLMIPMCKMLQDVFNPLVIQMT 123  
 QY 150 DDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKNVVKIQ 208  
 DB 124 DDEKFLFKQGVSLDDCCRFARENKDIIVAGFDPKTFIFMNSTYVG--GAFYQNVVRIA 181  
 QY 209 KAVTFNQVKGIFGFTSDCIKISFPALQAPSPNSFPQIFRRTDIOCLIPCAIDDP 268  
 DB 182 KCTIANSKACFGFTSDSISGKIHFAIOAPSSSPFPHFNGAKDIPCLIPCAIDDP 241  
 QY 269 YFMTDVAAPRIGYPRKALHSTFEPALQAGQTMASDPNSIFLTDPAKQIKTKVYK 328  
 DB 242 YFHLTRVSGRLKPKKALHSHRFPALQGPQSMASKOSALFMTDTNKKIKNKNRH 301  
 QY 329 AFSGGRDTEIEHROFGNCVDVSPMYLTFLEDDKLEQIRKDYTGAMLTGELKALI 388  
 DB 302 AFSGGQTIIEHREKGNPDVAVQYLSFLDDDEKLEQIYNTYKAGTLSTGEMKECI 361  
 QY 389 EVLQPLAEHQARKEVTEIIVKEFMT-PRKLSF 421  
 DB 362 KLLQGFVSDFOARSKVDKATLDMFMDSKLEW 395

## RESULT 6

S51901

tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein HR432; protein O0792; tryptophanyl-tRNA synthetase

C/Species: *Saccharomyces cerevisiae*

C/Date: 05-May-1995 #sequence revision 03-Aug-1995 #text\_change 03-Jun-2002

A/Accession: S51901; S59177; S66793

R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.

submitted to the EMBL Data Library, January 1995

A/Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including and a Delta.

A/Reference number: S51848

A/Accession: S51901

A/Molecule type: DNA

A/Residues: 1-432 &lt;VAM&gt;

A/Cross-references: EMBL:Z48149; NID:9663234; PIDN:CAA88164.1; PID:9663256

R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.

Yeast 11, 1069-1075, 1995

A/Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the delta element.

A/Reference number: S59156; MUID:96076631; PMID:7502582

A/Accession: S59177

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-432 &lt;VAM&gt;

A/Cross-references: EMBL:Z48149; NID:9663234; PIDN:CAA88164.1; PID:9663256

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

R/Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S66791

A/Accession: S66793

A/Molecule type: DNA

A/Residues: 1-432 &lt;DUR&gt;

A/Cross-references: EMBL:Z74839; NID:91419947; PIDN:CAA91110.1; PID:91419948; GSPDB:GN00

A/Experimental source: strain S286C

C/Genetics:

A/Gene: SGD:WRS1; WRS1; MIPS:YOL097C

A/Cross-references: SGD:S0005457

A:Map position: 15L

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C/Key word: ligase

F:117-120/Region: ATP-binding motif (HXGH)

Query Match 51.8%; Score 1163; DB 2; Length 432;

Best Local Similarity 54.8%; Pred. No. 1.1e-86;

Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 30 ATEAEDFVDPWTV-----QTSAGKIDYDKLIVRGSSKIDKELINRERATGGRPHF 84  
 DB 19 STDVKEQVATPMVEGEGVDEGRAGNIDYDKLQKGTGRVNEETLKRFQVGRPREPHF 78  
 QY 85 LRGRIFPSHRDNNQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFITKMLQDVFNPL 144  
 DB 79 LRGLFSEEDPFKILDLVBQGFPLLYTGRGSSSMHLMIPMCKMLQDVFNPL 138  
 QY 145 VIQMTDDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKN 203  
 DB 139 VIELTDEKFLFKHKLTIQVKNFAENKDIIVAGFDPKNTFIFSDLYMG--GAFYET 196  
 QY 204 VKIQKHVTENQYKIFGFTSDCIKISFPALQAPSPNSFPQIFRRTDIOCLIPCA 263  
 DB 197 VVRSQITGSTAKAVFGFNDSDCIKGFHFAIOAPSSSPFPHFNGAKDIPCLIPCA 256  
 QY 264 IDDPYFMTDVAAPRIGYPRKALHSTFEPALQAGQTMASDPNSIFLTDPAKQIKTK 323  
 DB 257 IDDPYFRCRDVADLTKRSKPLHSHRFPALQSTTMASDPTTALFMTDTNKKIKN 316  
 QY 324 KVKAKHAFSGGRDTEIEHROFGNCVDVSPMYLTFLEDDKLEQIRKDYTGAMLTGEL 383  
 DB 317 KINKVAFSGGQVSADLHRELGNPDVAVQYLSFFKDDVFLKECYDKYKSGELLSEW 376  
 QY 384 KKALIEVLQPLAEHQARKEVTEIIVKEFMT-PRKLSF 419  
 DB 377 KKLCTIETLOEFVAKFQERRAOVDEETLDKFMVPHKL 412

## RESULT 7

C90190

tryptophanyl-tRNA synthetase (trps) [imported] - *Sulfolobus solfataricus*C/Species: *Sulfolobus solfataricus*

C/Date: 24-May-2001 #sequence revision 24-May-2001 #text\_change 15-Jun-2001

A/Accession: C90190

R/Singh, R.K.; Cofaloni, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-Y

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: C90190

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-386 &lt;KUR&gt;

A/Cross-references: GB:AE00641; NID:913813608; PIDN:AAK40778.1; GSPDB:GN00155

C/Genetics:

A:Gene: trps

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 40.4%; Score 907; DB 2; Length 386;

Best Local Similarity 48.6%; Pred. No. 6.2e-66;

Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

QY 31 TEAEDFVDPWVQTSANG-IDYDKLIVRGSSKIDKELINRERATGGRPHFARGI 89  
 DB 6 TMDDEFTVTPWEV-----KGKVDYDKLIVQFGTKTEELKORIKNLAIDL-HVMLRRNV 59  
 QY 90 FFSHRDNNQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFITKMLQDVFNPLVIQMT 149  
 DB 60 FFSHRDLDLVLDYKSGKPFYLYTGRAPSL-GMHLGLLPFTKMLQKFNANLVIET 118  
 QY 150 DDEKYLMD-LTLDOAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKNVVKIQ 208

Db 119 DDEKYMNEPFTLQTRSMAYDNILDLIAVGFNDFPIFDQTEYL---RNMVPIYKIA 175  
 QY 209 KAHFENQYKGFEGTDDDCIGKISFPALQAPBSNSNRPQIFRFRRTIOCLIPALIDOP 266  
 Db 176 KKLTFSEKRAFGLDASNSIGLIFYPALQIAFT-----WFEKK---RCLIPAGIDOP 225  
 QY 269 YFRMTRDVAIRIGYKPKPALIHSTFPPALQAGQTMASDPNSSIFLTDPAKQIKTKVNH 328  
 Db 226 YMRJORDIASLGYKKAQIHSKELPILTGEGKSSSNPEATAIYLVDDPVTYARKIMKY 285  
 QY 329 AFSGGARDTIEHROFGNCVDVGSFMYLTFLIEDDD--KLEJIRKQYNSGAMLTGELKAL 387  
 Db 286 AFSGGQPTIELHRKYGNGNPEIDVPFWMLYFPEEDDNRKIEEHSYNSGKMLTGELAQIL 345  
 QY 388 IEVJQPLIAEHQARKEVTDIEIKPEMTRKLS 420  
 Db 346 IDKLNNLEBHR--RRREBAKELVAVFYKDGKIA 377

RESULT 8  
C75020  
tryptophanyl-tRNA synthetase (trps) PAB1111 - *Pyrococcus abyssi* (strain Orsay)  
C/Specties: *Pyrococcus abyssi*  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C/Accession: C75020  
R/anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A/Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure  
A/Reference number: A75001  
A/Accession: C75020  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-385 <KAM>  
A/Cross-references: GB:AU246288; GB:AL096836; NID:G5458960; PIDN:CBMS0601.1; PID:G5459111  
A/Experimental source: strain Orsay  
C/Genetics:  
A/Gene: trps; PAB1111  
A/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

[illegible]

RESULT 9  
 G71206  
 t-tryophan-tRNA ligase (EC 6.1.1.2) - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 03-Jun-2002  
 C:Accession: G71206  
 R:Kawarabayashi, Y.; Sawade, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekino, M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: G71206  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-301 #KAM-  
 A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BA031046.1; PID:g3258363  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank C:Genetics:  
 A:Gene: PH1921  
 C:Superfamily: yeast tyrosine-tRNA ligase  
 C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match	27.9%	Score 626.5;	DB 2;	Length 301;
Best Local Similarity	44.2%;	Pred. No. 2.8e-43;		
Matches 134;	Conservative 56;	Mismatches 92;	Indels 21;	Gaps 7;

122 MHVGHLLPFIIFKMWODVFNVPILVIOMTDDKYLAKD-LTLDAQYGAVERAKDIIACGF 180  
 Db 1 MHGIIIFPAIKMLQEKGVNLVYIOTIDDEKTLFKENLTFDPTKKRAVYNIIDIIANGF 60

RESULT 10  
G84373  
tryptophanyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1  
C1Species: Halobacterium sp. NRC-1  
C1Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C1Accession: G84373  
R1Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
J.; Letshauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A1Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A1Title: Genome sequence of Halobacterium species NRC-1.  
A1Reference number: A84160; M0ID:205044B3; F0ID:11016950  
A1Accession: G84373  
A1Status: preliminary  
A1Molecule type: DNA  
A1Residues: 1-380 <STO-  
A1Cross-references: GB:AB004437; NID:gi10581646; PIDN:AAG20355.1; GSPDB:GN00138  
C1Genetics:  
C1Gene: trpS2  
C1Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 23.8%; Score 534.5; DB 2; Length 380;  
 Best Local Similarity 35.4%; Pred. No. 1.2e-35;  
 Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;  
 QY AEDPVDPTVOTSSAKIDYDKLIVRGSSKIDKELINIRATGQRPHHFLRGIFPS 92  
 Db ADGNDVTPAVESDD---LDYEKILARFAGDELTDGRAPF-----DHLVNRGLFEYA 53  
 QY 33 HRDMNOVLDAVENKPPFLYITGGRSSSEAMHVGHLIFIFTKMLQDVNVPLVYQMTDDE 152  
 Db 54 GRVDVDFLTAGEOS---IYTVGVPSSG-PHMLGHAMVFYARRIQDFGARVYVPLSDDE 108  
 QY 153 KYLMKDLTLLOAAGDAVE-NAKDIIACGPDINKTFIF---SDLDYV-GMSSGFYKXVVKI 207  
 Db 109 KYMFKQOTPAET-GDYIRALRLDLAVGFPELRIYVDRDADVLVPLATAPAGDV--- 164  
 QY 208 QKAVTFNQVKGIFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCALDQ 267  
 Db 165 -RHATLQNVG-----EPDNYGQAFYPAVQTAHL---LPQLVHG--EHETLVPIAVDQD 213  
 QY 268 PYRRMTDVAPRIGYP--KRALHSTFPALQAGQOTMSASDPNSSIFLTDTAQKITKY 325  
 Db 214 PHVRVSDVAAKARYPVGKFGALMQLPFLAG-PGMSSS-AGVSIKLTDSPTVAREKV 271  
 QY 326 NKAFSGGRDTIEHROFGNCVDVDFMYLTFLEDD-D-KLBOIRKDYSGAMLTGELK 384  
 Db 272 RTIAYTGSRASVEHRKAGVPADVPFOYLSAFEDDDELARIERYAGDLSGELK 331  
 QY 385 KALIEVLQPLIAEHQARRKEVTEIYKEMTPRKLSFD 422  
 Db 332 DLAAADRIEFLAAGRRALGD--VTEALDAFRLTD 367

## RESULT 11

664476  
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Methanococcus jannaschii  
 C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 03-Jun-2002  
 C/Accession: F64476  
 R/Built: C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 Ison, J.D.; Sedow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurest, M.A.  
 Science 273, 1058-1073, 1996  
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C.  
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A/Reference number: A64300; MUID:96337999; PMID:8688087  
 A/Accession: F64476  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-370 <BUIL>  
 A/Cross-references: GB:U67582; GB:L77117; NID:g1592064; PIDN:AAB99425.1; PID:g1592065;  
 C/Genetics:  
 A/Map position: FOR1375885-1376997  
 A/Start codon: GTG  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.2%; Score 409.5; DB 2; Length 370;  
 Best Local Similarity 30.9%; Pred. No. 1.7e-25;  
 Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
 QY 40 PWTVOSSAGIDYDKLIVRGSSKIDKELINIRATGQRPHHFLRGIFSHRDMNV 99  
 Db 8 PW--ETPAV--IDYKTMQFGVAKPIVDVGLDEK-----HFFFRNIIIGHRDERI 57  
 QY 100 LDAYENKPPFLYITGGRSSSEAMHVGHLIFIFTKMLQ--DVFNVPVLVYQMTDDEKYLW 156  
 Db 58 VDAIKNNKEFAVVSQMMPSGR-WHFGKMYVVDLLKFOKXTDNINIFI-----ADLEAYMA 112  
 QY 157 KDILTDQAYDAV-ENAKDIIACGPDINKTFIFSDLDYMGSSGFYKXVVKI--OKAVTEN 214

Db 113 RNMSFETKELALNXYITNYIALGLDPEKINYYLQSKQGV-----KDLALILSKRTNMS 167  
 QY 215 QVAGIGFTDSDCIKISFPALQAPSPNSFPQIFRDT--DIQCLIPATIDDDYFEM 272  
 Db 168 EMKATYGFGETNIGHVFAPIYQVADIL---HQDLNLSPEPKVPVAVPGIDODPHIRL 224  
 QY 273 TRDVAE---IGYKRALHSTFPALQAGQOTMSASDPNSSIFLTDTAQKITKYKNA 329  
 Db 225 TRDIANRAKEFKRIPPSSTYHRMTGLLG--KMSSKRETAFLTDDETKVKKIFS-A 281  
 QY 330 FSGGRDTIEHROFGG--NCDVDSFMYLTFLEDDDKLBOIRKDYSGAMLTGELKKA 386  
 Db 282 KTGGRTELEHKKYGVPEECVYELFLY--HLILDKEIAETIYQCRSGELTGCKCKM 339  
 QY 387 LIEVLQPLIAEHQARRKEVTEIYK 411  
 Db 340 AYERVVEFLDKLKEKREKAEIAVK 364

## RESULT 12

669131  
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Methanobacterium thermoautotrophicum  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Jun-2002  
 C/Accession: E69131  
 R/Smith, D.R.; Doncelte-Stamm, L.A.; DeJonghe, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ujwani, N.;  
 Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi  
 A/Reference number: A69000; MUID:98037514; PMID:9371463  
 A/Accession: E69131  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-364 <MTH>  
 A/Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB4757.1; PID:g2621301  
 A/Experimental source: strain Delta H  
 C/Genetics:  
 A/Gene: MTH251  
 A/Start codon: TTG  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 17.7%; Score 397.5; DB 2; Length 364;  
 Best Local Similarity 27.9%; Pred. No. 1.6e-24;  
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;  
 QY 38 VDFWVOTSSAKIDYDKLIVRGSSKIDKELINIRATGQRPHHFLRGIFSHRDMN 97  
 Db 2 IDPW-----GSAK-LEVDLIENGVAPF-SEVLDEV-----PEPSWLMRRGIIFGHRDYE 50  
 QY 98 QUIDAYENKPPFLYITGGRSSSEAMHVGHLIFIFTKMLQDVNVPLVYQMTDDEKYLW 157  
 Db 51 RIISAMKGEDFAVYVGMMPSSGR-MAIGHKMYVDQLRW-YDRMGABIFIPIDAMEYVSAR 108  
 QY 158 DLTLDQAYGDAVEN-AKDIIACGPDINK-----TFIFSDLDYMGSSGFYKXVVK 206  
 Db 109 GUDFESRRIALIEYIAGTIALGLDEKONIHVYLOSEMLWEDLAVY----- 156  
 QY 207 IQKAVTFNQVKGIFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCALDQ 266  
 Db 157 LAGKVNFELRAIYGTGSTMHMYAPLIQVSDILHPLQDELGGR---PVLVPGPDQ 213  
 QY 267 DPFYRMTDVAARI---GYKRALHSTFPALQAGQOTMSASDPNSSIFLTDTAQKIT 322  
 Db 214 DPHIRLIRDIARFRDRYFILLSSSTYHRFMGLTGG--KMSNRRKSAIFLSDTEEAR 271  
 QY 323 TKYNKAFSGGRDTIEHROFGNCVDVDFMYLTFLE--DDDKLEIRKDYSGAMLTG 381  
 Db 272 AKI-RNAKGRGRTLEKQRELGCVPECCITETILYHMGSSBSRLBIYESCRNGTLMG 330  
 QY 382 ELKALIEVLQPLIAEHQARRKE 404

Db 331 ECKNNNAEPIRKFEELSVKREK 353

RESULT 13

T43806

tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)

C.Species: Encephalitozoon cuniculi

C.Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 03-Jun-2002

C.Accession: T43806

R.Peyretallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.

Mol. Biol. Evol. 15, 683-689, 1998

A.Title: Microsporidia, amitochochordal protists, possess a 70-kDa heat shock protein gen

A.Reference number: Z22693; MUID:98277683; PMID:9615449

A.Accession: T43806

A>Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-134 <PEY>

A.Cross-references: EMBL:AJ012470; PIDN:CAA10034.1

C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C.Keywords: ligase

Query Match 17.2%; Score 386; DB 2; Length 134;

Best Local Similarity 51.5%; Pred. No. 3.2e-24;

Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 35 EEPVDMVYQTSASAK----GIDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGIF 90

Db 3 EORITPMDVEVSTGEVPAIDYDKLINGCEKFNQALADRLKLSGRPAHYFFRKGIV 62

QY 91 FSHRDMNOVLDAVENKKPFLYLTGGRGSEAMHGHILPFIPTKMLQDVENVPLVYQMTD 150

Db 63 FHRDNLNLDLRIANNRPFLYLTGGRGSKTMHIGITIFPLCKYQDAFKIRLVYQITD 122

QY 151 DEKYLMDKLTLD 162

Db 123 DEKFLMKSWRL 134

RESULT 14

E69461

tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C.Species: Archaeoglobus fulgidus

C.Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C.Accession: E69461

R.Klienk, H.P.; Claydon, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A.Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A.Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A.Reference number: A69250; MUID:98049343; PMID:9389475

A.Accession: E69461

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-420 <KLE>

A.Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PIDN:AA8954.1; PID:9264885

C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 3e-22;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 38 VDPMTVQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMN 97

Db 3 VDPMEVEGV----IDSKLIIEFGMQPF--SEVLPEID-----NPHILMRGALIFGRDYW 52

QY 98 QVLDAVENKKPFLYLTGGRGSEAMHGHILPFIPTKMLQDVENVPLVYQMTDEKYLTK 157

Db 53 RLIEAMQKEPMAVWSGFMSG--LPHFGHKMTDELIVWQSAGKAFV--AIDMEASHVA 110

QY 158 DLTLDAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKNV--KIQKHVTFNQV 216

Db 111 GLSMKTEBELGMYLKSIIALGLREDATYYGQ-----KSHYVDLAEFLSAFVNFSEL 164

QY 217 KGIFFPTSDICIGKISFPALQAPSPFSNSFPQIFPRDITDQCLICALIDDDYFPMTRDV 276

Db 165 RAIVGFNSDTSIAKMFVTAIQADIL---HPQISDFGPKRVVVVGVGADDDPHMSLTDL 221

QY 277 APR----- 280

Db 222 AARISFSPFVEGVRVRSRKAELYSGLDLDFDKKIYEEHMDIFGEAEIERAVRKI 281

QY 281 -----GYPKALHSTFFPALQAGTQKMSADPNSSIFLTDYAKQITKVNKIAFSGG 333

Db 282 EVELGFAFAPISPSSTYHFTTGLTGG--KMSSKRESYISLIDPREBAKMYMK--AFYGG 338

QY 334 RDTIEHROFGNCVDVDSFMYLTFLELD--DDKLEQIRKDYTSGLMTGELKALIEVLQ 392

Db 339 RATAEORRLTGGEPDRCVFEVLYSFLTIDSDLELNQIEACREGRLCGCKMAAEIVK 398

QY 393 PLIAEHQARKEV 405

Db 399 SFLKEHQEKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)

C.Species: Aeropyrum pernix

C.Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C.Accession: D72477

R.Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hakiwara, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A.Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyru

A.Reference number: A72450; MUID:99310339; PMID:10382966

A.Accession: D72477

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-374 <KAN>

A.Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BA81476.1; PID:95106165

A.Experimental source: strain K1

C.Genetics:

A.Gene: APE2461

C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 15.7%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 6.7e-21;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 38 VDPMTVQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMN 97

Db 8 IDPW---GAVETIKDYDLRLTFFGIRPSEVL--PLLRKAGKERSFLMRGIIFFGRDFD 61

QY 98 QVLDAVENKKPFLYLTGGRGSEAMHGHILPFIPTKMLQ--DVFNVPVLYQMTDEKYLTM 156

Db 62 KLEKAKGGERAVAVLTGMPGSK--PHRHKLTVDOLYILQKNGFV--FVALADAEARAV 118

QY 157 KDLTLDAYGDAVEN--AKDIIACGFDINKT--FIFSDLDYMGSSGFYKNVVKIQKHVTFN 214

Db 119 RRIIGEEAVRIAAVEYIANNMIALGDPKOTERYFG---TVRGTGYFLILQIFSGKVTAA 174

QY 215 QVKGIFG--FTSDSDICIGKISFPALQAPSPFSNSFPQIFPRDITDQCLICALIDDDYFPM 273

Db 175 EMEALYVELTPAKMAASLT---QAADILHVQLDDEYGYR--HVVVPGADQDDHRLTL 227

QY 274 RDVAPR-----IGYPKALHSTFFPALQAGTQKMSADPNSSIFLTDYAKQITKVNKIA 329

Db 228 RDLADRMAGVVELERPASTYHFTTGLTGG--RKMSSKRESDTIFLTDPREBAKMYMK--A 284

QY 330 FSGGDTIEHROFGNCV--DVSNFVLYTFLEDDKLEQIRKDYTS---GAMLTGELK 385

Db 285 LTGGRATAEORRLTGGEPDRCVFEVLYSFLTIDSDLELNQIEACREGRLCGCKMAAEIVK 398

Fri Jul 11 11:50:10 2003

us-09-813-718-10\_copy\_48\_471.rpr

Page 8

Qy 386 ALIEVLOPLIAEHQARRKEVTDEIVKEFMTPR 417  
| : ||| : | : |  
Db 342 IAMEKLERFLAEHOSRLEKAKTIAMKLVBPFR 373

Search completed: July 10, 2003, 12:32:17  
Job time : 18.4461 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:26:29 ; Search time 9.25091 Seconds

(without alignments)  
1900.998 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEMPRKISPDQ 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_40:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	99.3	471	1	SYN_HUMAN
2	2134.5	95.0	475	1	SYN_BOVIN
3	2056	91.5	475	1	SYN_RABIT
4	2032	90.5	481	1	SYN_MOUSE
5	1210	53.9	395	1	SYN_SCHPO
6	1163	51.8	432	1	SYN_YEAST
7	910.5	40.5	381	1	SYN_SULTO
8	905	40.3	380	1	SYN_SULSO
9	803	35.8	385	1	SYN_PYRAB
10	796	35.4	386	1	SYN_PYRHO
11	735	32.7	375	1	SYN_PYRAE
12	534.5	22.8	380	1	SYN_HALNI
13	409.5	18.2	370	1	SYN_MERJA
14	397.5	17.7	364	1	SYN_MERTH
15	386	17.2	134	1	SYN_BNCCU
16	370.5	16.5	420	1	SYN_ARCFU
17	353	15.7	374	1	SYN_AERPE
18	299.5	13.3	426	1	SYN_THRVO
19	268.5	12.0	466	1	SYN_THRAC
20	192	8.5	323	1	SYN_ARCFU
21	185.5	8.3	341	1	SYN_STRPN
22	182.5	8.1	340	1	SYN_STRPY
23	175.5	7.8	341	1	SYN_CLOLO
24	174.5	7.8	341	1	SYN_LACLA
25	172	7.7	394	1	SYN_YEAST
26	170.5	7.6	366	1	SYN_SULSO
27	169	7.5	395	1	SYN_AQUAE
28	166.5	7.4	346	1	SYN_CHLTR
29	163.5	7.3	353	1	SYN_BOBBU
30	159.5	7.1	337	1	SYN_TREPA
31	158	7.0	346	1	SYN_CHLNU
32	152	6.8	344	1	SYN_CHLPN
33	146	6.5	326	1	SYN_HELPY

34	138	6.1	319	1	SYN_MERTH	027795 methanobact
35	136.5	6.1	328	1	SYN_BOVIN	029465 bos taurus
36	135.5	6.0	328	1	SYN_THRMA	090953 thermocoga
37	133	5.9	328	1	SYN_BACST	000953 bacillus st
38	132.5	5.9	343	1	SYN_MYCLE	049901 mycobacteri
39	132.5	5.9	528	1	SYN_HUMAN	P54577 homo sapien
40	130	5.8	326	1	SYN_HELPJ	092134 helicobacte
41	129	5.7	401	1	SYN_SCHPO	014055 schizosacch
42	128.5	5.7	379	1	SYN_YEAST	P04803 saccharomyc
43	128	5.7	334	1	SYN_HAEIN	P43835 haemophilus
44	124.5	5.5	347	1	SYN_MYCGE	P47372 mycoplasma
45	122	5.4	350	1	SYN_CAEEL	P46579 caenorhabdi

## ALIGNMENTS

RESULT 1	ID	SYN_HUMAN	STANDARD;	PRT;	471 AA.
AC	P23381	P78535; Q9UDJ3;			
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-DEC-1992	(Rel. 24, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)				
DE	(TrpRS) (TFP53) (hmRS).				
GN	WARS.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92105071; PubMed=1761529;				
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;				
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human				
RT	fibroblasts.";				
RL	J. Biol. Chem. 266:24245-24248(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92107982; PubMed=1763065;				
RA	Flechner J., Rasmussen H.H., Justesen J.;				
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa				
RT	protein (gamma 2) highly homologous to rabbit peptide chain release				
RT	factor and bovine tryptophanyl-tRNA synthetase.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92112058; PubMed=1765274;				
RA	Prolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovjeva O.L.,				
RT	Kiselev L.L.;				
RT	"Cloning and nucleotide sequence of the structural gene encoding for				
RT	human tryptophanyl-tRNA synthetase.";				
RL	Gene 109:291-296(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92164636; PubMed=1537332;				
RA	Buhtic U., Flohr T., Boettger E.C.;				
RT	"Molecular cloning and characterization of an interferon induced				
RT	human cDNA with sequence homology to a mammalian peptide chain				
RT	release factor.";				
RL	EMBO J. 11:485-496(1992).				
RN	[5]				
RP	SEQUENCE OF 1-13 FROM N.A.				
RX	MEDLINE=96319994; PubMed=8724762;				
RA	Sokolova I.V., Narovjanskii A.N., Amchenkova A.M., Turpaev K.T.;				
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-				
RT	tRNA synthetase gene.";				
RL	Mol. Biol. (Mosk) 30:319-329(1996).				
RN	[6]				
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.				
RC	TISSUE=Spem;				
RX	MEDLINE=93292992; PubMed=7685728;				

RA Frolova L.Y., Grigorieva A.Y., Sudomocina M.A., Kisselev L.L.;  
 RT "The human gene encoding cryptophanyl-tRNA synthetase: interferon-  
 RL response elements and exon-intron organization.";  
 RN Gene 128:237-245 (1993).  
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.  
 RC TISSUE=Keratinoctyes;  
 RX MEDLINE=93162043; Pubmed=1286667;  
 RA Rasmussen H.H., van Damme J., Puyse M., Gesser B., Celis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RL protein database of normal human epidermal keratinocytes.";  
 RN Electrophoresis 13:960-969 (1992).  
 RP FUNCTION.  
 RX MEDLINE=92225128; Pubmed=1373391;  
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;  
 RT "An interferon-induced protein with release factor activity is a  
 RL cryptophanyl-tRNA synthetase.";  
 CC FBES Lett. 300:162-166 (1992).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- INDUCTION: BY INTERFERON GAMMA.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M77804; AAA67324.1; -  
 DR EMBL; X59892; CAA42545.1; -  
 DR EMBL; M61715; AAA61298.1; -  
 DR EMBL; X62570; CAA44450.1; -  
 DR EMBL; S82905; AAB39381.1; -  
 DR EMBL; X67920; CAB94198.1; -  
 DR EMBL; X67921; CAB94198.1; JOINED.  
 DR EMBL; X67922; CAB94198.1; JOINED.  
 DR EMBL; X67923; CAB94199.1; -  
 DR EMBL; X67924; CAB94199.1; JOINED.  
 DR EMBL; X67925; CAB94199.1; JOINED.  
 DR EMBL; X67926; CAB94199.1; JOINED.  
 DR EMBL; X67927; CAB94199.1; JOINED.  
 DR EMBL; X67928; CAB94199.1; JOINED.  
 DR PIR; A41706; A41706.  
 DR PIR; A41633; A41633.  
 DR PIR; JH0533; JH0533.  
 DR PIR; S19246; S19246.  
 DR Aarhus/Ghent-2DPAGE; 3524; IEF.  
 DR PHCI-2DPAGE; P23381; -  
 DR Genew; HGNC:12729; WARS.  
 DR MIM; 191050; -  
 DR InterPro: IPR002306; Trp tRNA-synt\_1b.  
 DR InterPro: IPR000738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR Pfam: PF00458; WHEP-TRS; 1.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRAME; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA tRNA\_LIGASE\_I; 1.  
 DR PROSITE; PS00762; WHEP-TRS; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT DOMAIN 19 64 WHEP-TRS.  
 FT SITE 164 173 "HIGH" REGION.  
 FT SITE 349 353 "KMSKS" REGION.  
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
 FT CONFLICT 424 424 A -> R (IN REF. 4).

SO SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;  
 Query Match 99.3%; Score 2231; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 6, 9e-177;  
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MSYKAAAGEDYKADCCPGNAPATSNHGPDAATEAEEDFPVDPWTQTSASAKIDVDKILVRF 60  
 DB 48 MSYKAAAGEDYKADCCPGNAPATSNHGPDAATEAEEDFPVDPWTQTSASAKIDVDKILVRF 107  
 QY 61 GSSKIDKEILNRIERATGQRPHPFLRGIFFSHRDMNOVLDAVENKPPFLYTGRCPSSE 120  
 DB 108 GSSKIDKEILNRIERATGQRPHPFLRGIFFSHRDMNOVLDAVENKPPFLYTGRCPSSE 167  
 QY 121 AMVGHILPIPIFKMLQDVFNVLVIQMTDDKYLKMDLTLDAYGDAVENADIIACGF 180  
 DB 168 AMVGHILPIPIFKMLQDVFNVLVIQMTDDKYLKMDLTLDAYGDAVENADIIACGF 227  
 QY 181 DIKRTIFSDLDYMGSSGPKVNVKIQKVPNOYKGIFFGFDSDICIGISPPAIQAP 240  
 DB 228 DIKRTIFSDLDYMGSSGPKVNVKIQKVPNOYKGIFFGFDSDICIGISPPAIQAP 287  
 QY 241 SFNSFPQIFRDRTDIQCILPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 300  
 DB 288 SFNSFPQIFRDRTDIQCILPCAIDDPYFRMTRDVAPRIGYKPKALLHSTFPALQAG 347  
 QY 301 TKMSASDPNMSITLTPTAQIKTKVKNKAFSGGRDTIEEHROGNCDDVDSFMYLTFFL 360  
 DB 348 TKMSASDPNMSITLTPTAQIKTKVKNKAFSGGRDTIEEHROGNCDDVDSFMYLTFFL 407  
 QY 361 EDDDKLEQIRKQVTSAMTGLKALIEVLOPLIAEHORARKEVDEIVKEEMTPRKLX 420  
 DB 408 EDDDKLEQIRKQVTSAMTGLKALIEVLOPLIAEHORARKEVDEIVKEEMTPRKLX 467  
 QY 421 FDFQ 424  
 DB 468 FDFQ 471  
 RESULT 2  
 ID SYM\_BOVIN STANDARD; PRT; 475 AA.  
 AC P17248;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpRS).  
 GN WARS.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina;  
 RX MEDLINE=91329348; Pubmed=1907847;  
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,  
 RA Gander J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,  
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;  
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 RT prokaryotic synthetases but near identity with mammalian peptide  
 RT chain release factor.";  
 RL Biochemistry 30:7809-7817 (1991).  
 RN [2]  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Garret M., Trezeguet V., Pajot B., Gander J.-C., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., Labouesse B.,  
 RA Labouesse J., Bonnet J.;  
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

```

CC diphasphate + L-tryptophanyl-tRNA (Trp).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53918; CAA37872.1; -
DR EMBL; X52113; CAA36356.1; -
DR PIR; A40279; YWBO.
DR PIR; S14540; S14540.
DR InterPro; IPR002306; Trp_cRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
KW DOMAIN
FT 1 69 WHEP-TRS.
FT 2 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
FT SITE 169 178 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT CONFLICT 17 17 L -> M (IN REF. 2).
SQ SEQUENCE 475 AA; 53729 MW; F7E531750137E332 CRC64;
Query Match 95.0%; Score 2134.5; DB 1; Length 475;
Best Local Similarity 95.0%; Pred. No. 6.7e-169;
Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;
2 SYKAAAGEYKADCPGPNAPTSNHCPRDTEAEEDVVDVWYVQTSAGKIDYDKLIVRG 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 SYKAAAGEYKADCPGPNAPTSNHCPRDTEAEEDVVDVWYVQTSAGKIDYDKLIVRG 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSA 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSA 173
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 MHVGHILPFIETKWLQDVFNPLVIOMTDDEKYLMDLTLDQAYGDAVENANDIACGPD 181
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
174 MHVGHILPFIETKWLQDVFNPLVIOMTDDEKYLMDLTLDQAYGDAVENAND-ITCGFD 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 INKTFPSLDYMGSSGFFKYKVVKIQKAVTENVQVGIFFGTFDSDCIGKISPAIOAPS 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 INKTFPSLDYMGSSGFFKYKVVKIQKAVTENVQVGIFFGTFDSDCIGKISPAIOAPS 292
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
242 FSNSPFOIRDRDRIQCLIPCAIDDPYFRMTRDVAPRIGYPPALHSTFPALQAGAT 301
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
293 FSNSPFOIRDRDRIQCLIPCAIDDPYFRMTRDVAPRIGYPPALHSTFPALQAGAT 352
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 KMSASDPNSSIFLTDPAKQIKTKVNGHAFSGGRTIEEHROFGNCDVDSFMYLTFPLE 361
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
353 KMSASDPNSSIFLTDPAKQIKTKVNGHAFSGGRTIEEHROFGNCDVDSFMYLTFPLE 412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 DDDKLEQIRKYTSGAMLTGELKKALIEVLOPIAHQARRKEVDEIYKEMTPPKLSF 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
413 DDDKLEQIRKYTSGAMLTGELKKALIEVLOPIAHQARRKEVDEIYKEMTPPKLSY 472
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
422 DFO 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 DFO 475
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 3

```

```

SW_RABIT
ID SW_RABIT STANDARD; PRT; 475 AA.
AC P25612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN MARS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigden W.J., Muzny D.M., Harlow E., Caskey C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RX MEDLINE=94009008; PubMed=8404867;
RA Prolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins.";
RL EMBL J. 12:4013-4019(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphasphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33460; AAA31246.1; ALT_SEQ.
DR EMBL; U02595; AAA60257.1; -.
DR PIR; A35904; YMRBPR.
DR InterPro; IPR002306; Trp_cRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
DR AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
KW DOMAIN
FT 1 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;
Query Match 91.5%; Score 2056; DB 1; Length 475;
Best Local Similarity 90.5%; Pred. No. 2.1e-162;
Matches 383; Conservative 20; Mismatches 20; Indels 0; Gaps 0;
2 SYKAAAGEYKADCPGPNAPTSNHCPRDTEAEEDVVDVWYVQTSAGKIDYDKLIVRG 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
53 SYKAAAGEYKADCPGPNAPTSNHCPRDTEAEEDVVDVWYVQTSAGKIDYDKLIVRG 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 SSKIDKELINRIERATGQRPHHFLRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSA 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 113 SSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 172
QY 122 MHVGHILPFIPTKMLQDVFNVLVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFD 181
Db 173 MHVGHILPFIPTKMLQDVFNVLVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFD 232
QY 182 INKTFISDLDYMGSSGFYKNNVKIKQHTFNQVKIFGTSDDCIKGISFPALQAPAS 241
Db 233 VNKTFISDLDYMGSSGFYKNNVKIKQHTFNQVKIFGTSDDCIKGISFPALQAPAS 292
QY 242 FSNSPFOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGQ 301
Db 293 FSNSPFOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGQ 352
QY 302 KMSASDPNSSIFLTDYAKQIKTKVKNKAFSGGRDTIEHROFGNCVDVDSFMYLTFPLE 361
Db 353 KMSASDPNSSIFLTDYAKQIKTKVKNKAFSGGRDTIEHROFGNCVDVDSFMYLTFPLE 412
QY 362 DDDKLEQIRKDYTGAMLTGELKKALLLEVLOPLAEHQAARKVETDEIVKEFMTPRKLSF 421
Db 413 DDDKLEQIRKDYTGAMLTGELKKALLLEVLOPLAEHQAARKVETDEIVKEFMTPRKLSF 472
QY 422 DFDQ 424
Db 473 HYQ 475 /

```

```

RESULT 4
SYM_MOUSE STANDARD; PRT; 481 AA.
ID _SYM_MOUSE
AC P32521;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: The short isoform is widely expressed, the
CC long form is found only in embryonic stem cells.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69656; CAA49347.1; -
CC EMBL; X69657; CAA49348.1; -
CC PIR; S31461; S31461.
CC PIR; S31462; S31462.
CC MGD; MGI:104630; WARS.
CC InterPro; IPR002306; Trp_tRNA-synt_1b.
CC InterPro; IPR000738; WHEP-TRS.

```

```

DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA Ligase_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPIC 476 481 MISSING (IN ISOBORN 2).
SQ SEQUENCE 481 AA; 54282 MW; B05N452C08074F52 CRC64;

Query Match 90.5%; Score 2032; DB 1; Length 481;
Best Local Similarity 89.6%; Pred. No. 2e-160;
Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSYKAAGEDYKADCEPPNAPATSNNGPDATSEEDFPDPMYVQTSSAKGIDYDKLIVRF 60
Db 52 MSYKAMGGEYKAGCEPPNAPATSNNGPDATSEEDFPDPMYVQTSSAKGIDYDKLIVRF 111
QY 61 GSSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 120
Db 112 GSSKIDKELINRERATGQRPHRFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGPSSA 171
QY 121 AMRVGHILPFIPTKMLQDVFNVLVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 180
Db 172 AMRVGHILPFIPTKMLQDVFNVLVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGF 231
QY 181 DIKKTIFSDLDYMGSSGFYKNNVKIKQHTFNQVKIFGTSDDCIKGISFPALQAPAS 240
Db 232 DIKKTIFSDLDYMGSSGFYKNNVKIKQHTFNQVKIFGTSDDCIKGISFPALQAPAS 291
QY 241 SFNSPFOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGQ 300
Db 292 SFNSPFOIFRDRDIOCLIPCAIDDDPYFRMTDVAIRIGYKPKALLHSTFPALQAGQ 351
QY 301 TKMSASDPNSSIFLTDYAKQIKTKVKNKAFSGGRDTIEHROFGNCVDVDSFMYLTFPLE 360
Db 352 TKMSASDPNSSIFLTDYAKQIKTKVKNKAFSGGRDTIEHROFGNCVDVDSFMYLTFPLE 411
QY 361 EDDDKLEQIRKDYTGAMLTGELKKALLLEVLOPLAEHQAARKVETDEIVKEFMTPRKLS 420
Db 412 EDDDKLEQIRKDYTGAMLTGELKKALLLEVLOPLAEHQAARKVETDEIVKEFMTPRKLS 471
QY 421 FDFQ 424
Db 472 FHFQ 475

RESULT 5
SYM_SCHPO STANDARD; PRT; 395 AA.
ID _SYM_SCHPO
AC 009692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TrpRS).
GN SPAC2F7.13C
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RX SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

```

[illegible]

```

Db      242  YFRLTRDVSGRLKTKKKKALLHSRFPALQPGQSGMSA KQSSAIFEMTDPRKIKNKINRH 302
Qy      329  AFGSGRDTIEHRRPGNCVDVDSFMTLTFLEDDDKLEQIRKYTSGAMLTGELKALI 388
      |||||
Db      302  AFGSGGATIEIHRKGNPNPDVAVAYQVLSFPLDDEKTKQLYNTYKAGTISGEMKGC 361
      |||||
Qy      389  EVLQPLAEHQARRKEVTEIDVKEFMT-PRCLASF 421
      ::|||
Db      362  KLLQGFVSDFOAARSKYDEATILDMFMQSGKLEW 395
      ::|||

RESULT 6
SYNC_YEAST
ID      SYNC_YEAST      STANDARD;      PRT;      432 AA.
AC      Q12109;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      TrypTopophanY1-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (TrypTophan--
DE      tRNA ligase) (TTPRS).
GN      WRS1 OR YOL097C OR HRE432.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX      NCBI_TaxID=4932;
FN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=96076631; PubMed=7502582;
RX      Vandenbol M., Durand P., Portetelle D., Hliger F.;
RT      "Sequence analysis of a 44 kb DNA fragment of Yeast chromosome XV
RT      including the Tyl-H3 retrotransposon, the sufl1(+) frameshift
RT      suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-Ia and a
RT      delta element.";
RL      Yeast 11:1069-1075 (1995).
RN      [2]
RP      FUNCTION.
RX      MEDLINE=97197969; PubMed=9046085;
RA      John T.R., Ghosh M., Johnson J.D.;
RT      "Identification and expression of the Saccharomyces cerevisiae
RT      cytoplasmic tryptophanyl-tRNA synthetase gene.";
RL      Yeast 13:37-41 (1997).
CC      -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC      diphosphate + L-tryptophanyl-tRNA(Trp).
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on ways
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; Z48149; CAAB8164.1; -.
DR      EMBL; Z74839; CAA99110.1; -.
DR      SGD; S0005457; WRS1.
DR      InterPro; IPR002306; TTP_tRNA-synt_1b.
DR      InterPro; IPR002305; tRNA-synt_1b.
DR      InterPro; IPR001412; tRNA-synt_1b.
DR      Pfam; PF00579; tRNA-synt_1b; 1.
DR      PRINTS; PRO1039; TRNASYNTTRP.
DR      TIGRFAMs; TIGR00233; tRPS; 1.
DR      PROSITE; PS00178; AA_tRNA_LIGASE_I; 1.
KM      Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT      SITE
FT      111
FT      120
FT      295
FT      299
FT      SITE
FT      432 AA; 49350 MM; CA08F169737E9736 CRC64;
SQ      SEQUENCE

Query Match      51.8%; Score 1163; DB 1; Length 432;
Best Local Similarity 54.8%; Pred. No. 1.2e-86;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

```

```

QY 30 ATEAEEDFVDPWTV-----QTSSAKGIDVCKLIVRGSSKIDKELINIRERATQGRPHH 84
DB 19 STVKEQVTPWPMVEGVDEQGRQNDYDKLTKQFETKVNBEETLKRQVGRREPHH 78
QY 85 LRAGIFFSHRDMQVLDAYENKKRPFYLYTGRGSSSEAMHVGLIPFIETKMLQDVFNVL 144
DB 79 LRKGLFSESEDFPKIIDLVEQGRPFYLYTGRGSSSDMHGMIIPFIETKMLQDVFNVL 138
QY 145 VIQMTDEKYLKML-DLTLDQAYGDAVENAKDIIACGFDINKTFIFSLDLYMGSSGYYK 203
DB 139 VIELTDEKELFKHKLITINVKNFARENAKDIIAVGDPKQTFIFSLQYMG--GAFYET 196
QY 204 VVNIQKHVTFNQVKGIFGFDSDICIGISFPALQAAFSFNSPQIFRDRTDIQCILPCA 263
DB 197 VVAVSQITSTAAVAVGFNDSDICIGFHFASIQIATAFPSSFPNVGLPDKTPCLPCA 256
QY 264 IDDDPYFRMTRDVAFRIGYKPKPALHSTFPALQAAQTKMSASDPNSIFLDTAKOIKT 323
DB 257 IDDDPYFRVGRDVADKTKYKSKPALHRSFPALQGSTTKMSASDPTAIEMTDIPKOIOK 316
QY 324 KVNKHAFFSGGRDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383
DB 317 KINKYAFSGGVASADLIRELGNDVAVAYOYLSFFKDDVFLKCYDYKSGELLSEGM 376
QY 384 KKALIEUPLIAEHQARKEVDEIVKEFMTPRKL 419
DB 377 KLCIEUPLIEFVKAFOERRAQVDEETLCKFVPHKL 412

RESULT 7
SYM_SULTO
ID _SYM_SULTO STANDARD; PRT; 381 AA.
AC Q976MT;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpS)
GN TRPS OR ST0169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; Pubmed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Ohima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RT DNA Res. 8:123-140(2001)
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
RL -1- diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AP000981; BAB65126.1; ALT_INIT.
CC DR Interpro; IPR002306; Trp_tRNA-synt_1b.
CC DR Interpro; IPR002305; tRNA-synt_1b.

```

```

DR Pfam; PF00579; tRNA-synt_1b; 1.
DR TRIGRAME; TRIGR00233; trps; 1.
DR PROSITE; PS00178; AA_tRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 82 90 "HIGH" REGION.
FT SITE 254 258 "KMSKS" REGION.
SQ SEQUENCE 381 AA; 44718 MW; 109F5A56AFD7159 CRC64;

Query Match 40.5%; Score 910.5; DB 1; Length 381;
Best Local Similarity 50.5%; Pred. No. 7,9e-68;
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

QY 35 EDF-VDPWTVQSSAKG-IDYDKLIVRGSSKIDKELINIRERATQGRPHHFRAGIFPS 92
DB 3 QDFNVPMVEY-----KGKVDYDLVQFGTQKITSLSKKKISINDELHWMKRVFSS 57
QY 93 HRDMQVLDAYENKKRPFYLYTGRGSSSEAMHVGLIPFIETKMLQDVFNVLQMTDDE 152
DB 58 HRDLVLDVMDYQDGKGFPLYTGRAPSL-GMHLGHLIPFIETKMLQDKFNVNLVIEITDE 116
QY 153 KYIWK-DLTLDQAYGDAVENAKDIIACGFDINKTFIFSLDLYMGSSGYYKVVKIQKV 211
DB 117 KFMENPEYTLQTRQWADYNDIIDLIVGFNPDKTFIFQDTEYI--RMWPIAIKIAKFL 173
QY 212 TFNQVKGIFGFDSDICIGISFPALQAAFSFNSPQIFRDRTDIQCILPCAIDDPYFR 271
DB 174 TRSEVATGTLDTSSNIGIITWPAIQIAFT-----WEKR---RCLIPAGIDDPYFR 223
QY 272 MTRDVAFRIGYKPKPALHSTFPALQAAQTKMSASDPNSIFLDTAKOIKTKVNKHAFFS 331
DB 224 LQRDIAESIGYVQAQIHSKFLPPLTGPBGKSSSQPETAIVLTDPKVERKIMKAFSS 283
QY 332 GGRDTIEHRQFGNCDVVSFMYLTFPLE-DDDKLEQIRKDYTSGLMTGELKKALIEV 390
DB 284 GGQPTIELHRKKGNDIVDSFQWLYMFEPEDNKIKLIEDYRSGLLTGELKQTLIEK 343
QY 391 LQPLIAHQARKE 404
DB 344 LNDPFLERHOKKE 357

RESULT 8
SYM_SULTO
ID _SYM_SULTO STANDARD; PRT; 380 AA.
AC Q97ZX0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpS)
GN TRPS OR SS00452.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332256; Pubmed=11427726;
RA She Q., Singh R.K., Contaloni F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Toletup N.,
RA Charlebois R.L., Doolittle W.F., Duguet W., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
RL -1- diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.

```

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: A006677; AAK40778.1; ALT. INT.  
 DR InterPro: IPR002306; Trp. tRNA-synt. 1b.  
 DR InterPro: IPR002305; tRNA-synt. 1b.  
 DR Pfam: PF00579; tRNA-synt. 1b; 1.  
 DR PRINTS: PRO1039; TRNASYNTHTRP.  
 DR TIGRPFAM: TIGR00233; trps; 1.  
 DR PROSITE: PS00178; AA. TRNA. LIGASE I; FALSE. NEG.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 81 89 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 380 AA; 44691 MW; CF8344CF63883680 CRC64;  
 Query Match 40.3%; Score 905; DB 1; Length 380;  
 Best Local Similarity 48.7%; Pred. No. 2.2e-67;  
 Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;  
 QY 34 EEDFVDPWVYQTSASAG-IDYDKLIVRFSGSKIDKELINRIERATGQRPHHFLRGIFFSH 92  
 DB 3 DEFTVPMVEV-----KQKVDYDKLIVQFTQKITEELKORIKVLAIDL-HVMLERRVVFPS 56  
 QY 93 HRDMNOVLDAVENKKEFFLYLTGRGSSSEAMVGHLLPFITFKMLQOVENVPLVIQMTDDE 152  
 DB 57 HRDLDLVANDYKSKFFLYLTGRAPSL-GMHCHLIPFITFKLOKFRNALYITETDDE 115  
 QY 153 KILWK-DLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVKIQKEV 211  
 DB 116 KVMRNEFEFTLDQRSWAVYNIILDIIVGFNDPKTFIFQDEYI---RNMVPIVTKIAKCL 172  
 QY 212 TENOVGIGFTDSDICIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFR 271  
 DB 173 TSEVAVATGGLDASSNIGLIFPALQIAPL-----MEKK---RCLIPAGIDQDPYR 222  
 QY 272 MTRDVAIRIGYPPKALLHSTFFPALQAOAKSPNSIFLDTAKOIKTUVNCHAPS 331  
 DB 223 LQDIDIESIGYVYAAQIHSEKPLPLTGPEGKSSNPERAIYIVDPKVERKIMKTAAS 282  
 QY 332 GGRDTIEHRQFGNCDDVVSFMYLTFPLEDD-KLEQIRKDYTSAMLGELKALIEV 390  
 DB 283 GGQPIELHRKYGNEIDVDFQWLYYFPEEDNRIKEIEEYRSKMLTGLKQILIDK 342  
 QY 391 LQPLIAEHQARKKEVTDVKEFMRPKIS 420  
 DB 343 LNNFLEHR-RRREAKELVHVKYDGKLA 371  
 RESULT 9  
 ID SYM PYRAB STANDARD; PRT; 385 AA.  
 AC Q9UY11;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpRS).  
 GN TRPS OR PAB1111.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_Taxid=29292;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G5 / Orsay;  
 RA Heilig R.;

RT Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
 RT structure and evolution";  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBD databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC dihydrophate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ248288; CAB50601.1; -  
 DR InterPro: IPR002306; Trp. tRNA-synt. 1b.  
 DR InterPro: IPR002305; tRNA-synt. 1b.  
 DR InterPro: IPR001412; tRNA-synt. 1.  
 DR Pfam: PF00579; tRNA-synt. 1b; 1.  
 DR PRINTS: PRO1039; TRNASYNTHTRP.  
 DR TIGRPFAM: TIGR00233; trps; 1.  
 DR PROSITE: PS00178; AA. TRNA. LIGASE I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 82 90 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 385 AA; 45100 MW; 4C29D0141976B12 CRC64;  
 Query Match 35.8%; Score 803; DB 1; Length 385;  
 Best Local Similarity 45.3%; Pred. No. 6.2e-59;  
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;  
 QY 35 EDF-VDPWVYQTSASAGIDYDKLIVRFSGSKIDKELINRIERATGQRPHHFLRGIFFSH 93  
 DB 3 EDFKVTPEVGEV-----VDYNNLIHFGSPLTELEKTAELTSELPLFFRRKFFFSH 58  
 QY 94 RDMNOVLDAVENKKEFFLYLTGRGSSSEAMVGHLLPFITFKMLQOVENVPLVIQMTDDE 153  
 DB 59 RYDVKVLQDYERGRFFLYLTGRGSPG-PMHIGHIIPFATKMLQKFGVNLVLIQTIDDEK 117  
 QY 154 YLMKD-LTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKVVKIQKHT 212  
 DB 118 FLFKENLTFEDTKHVAVENIILDIIVGPDPKTFIFQNSEF---TKYEMAIPIAKKIN 173  
 QY 213 FNOVGIGFTDSDICIGKISFPALQAPSFNSFPQIFRDRDIOCLIPCAIDQDPYFRM 272  
 DB 174 FEMAAVAVGFTIEQSNIGMIFPALQIAPL-----FEKR---RCLIPALDQDPYRRL 223  
 QY 273 TRDVAIRIGYPPKALLHSTFFPALQAOAKSPNSIFLDTAKOIKTUVNCHAPS 332  
 DB 224 QRDFAESIGYVYAAQIHSEKPLPLTGPEGKSSNPERAIYIVDPKVERKIMKTAAS 283  
 QY 333 GGRDTIEHRQFGNCDDVVSFMYLTFPLEDDKLEQIRKDY---TSGAMLGELKALIE 389  
 DB 284 GRPITLKEQREKGEPEKCVFEMLEIFPEBDK-KLERYIACKXNGELTGCECKRYLS 341  
 QY 390 VLQPLIAEHQARKKEVTDVKEFMRPKIS 411  
 DB 342 KIQEFLKEHQKRRKKAQKQIEK 363  
 RESULT 10  
 ID SYM PYRHO STANDARD; PRT; 386 AA.  
 AC A05584;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpRS).  
 GN TRPS OR PH1921.

OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kousugi H., Hosoyama A., Nagai Y.,  
 RA Sekai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,  
 RA Masuchi Y., Shitaya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Ref. 5:55-76(1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AP000007; BAA31046.1; ALT: INIT.  
 DR InterPro: IPR002306; Trp-tRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA-TRNA-LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KM Complete proteome.  
 FT SITE 82 90 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 386 AA; 45305 MW; 9E3C39284028B2D CRC64;  
 Query Match 35.4%; Score 796; DB 1; Length 386;  
 Best Local Similarity 44.0%; Pred. No. 2,4e-58;  
 Matches 172; Conservative 68; Mismatches 125; Indels 26; Gaps 9;  
 QY 34 EEDFVDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFPSH 93  
 DB 3 EEFRTVPEVEGV-----VDYDKLKHGTSPLTEDELEKTAELTKSELPIFFRRKFFPSH 58  
 QY 94 RDMNQVLDAYENKKPFYLYTGRGSSSEAMHGHLPFTFKWLDVFNVLVQMTDDEK 153  
 DB 59 RVDLDLILKQVEBGRGFLYTGRRGSG--PMHGHLPFAFKMLOEKRGVLYLQITDDEK 117  
 QY 154 YLWKD-LTLDQAYGDAVENAKDIACGFDINKTFIPSDLDYMGSSGFFYKVVVKIQKHT 212  
 DB 118 FLTKENLTFTDTRKMAWDNLDIIVAGFDPDKTFIFQNSSEF---TKIYEMALPIAKIN 173  
 QY 213 FNVKVGIFGFTSDDCIGKISFPALQAPSPNSFPQIFRRTDIQCLIPALIDDPFERM 272  
 DB 174 FSNAAKVFGEESKSGIMIFFPALQAPTF-----FEERK---RCLIPALIDDPFERM 223  
 QY 273 TRDVAPRIGYKPKALHSTFPALQAGTQKMSADPNSSIFLNDTAKQITKTKNKAHPSG 332  
 DB 224 QRFASLSGYYKTAALHSKVPISLISGKMSKSPETALYILDSPEDEVAKYWKVFTLTG 283  
 QY 333 GRDTIEBHPGFGNCDDVDSFMVLTFFLEDDDKLEQIRKDY---TSGAMLTGELKALIE 389  
 DB 284 GRPTLKEQRKGEPEKCVVFKWLEIFFEEDDK--KLKERYVACRKGELTGCECKRYLIS 341  
 QY 390 VLOPLIAEHQARKKEVDEIVKKEFMPRKLS 420  
 DB 342 KIOEFLKEHORRRK--AEKLVKFKYTKGLA 371

RESULT 11  
 SYM PYRAE  
 ID SYM PYRAE STANDARD; PRT; 375 AA.  
 AC Q82T05;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (trpS).  
 GN TRPS OR PAE3091.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX PubMed=11792869;  
 RA Fitz-Gibbon S.T., Lader H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE009911; AL64664.1; -  
 DR InterPro: IPR002306; Trp-tRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1b.  
 DR TIGRFAMs: TIGR00233; trpS; 1.  
 DR PROSITE: PS00178; AA-TRNA-LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KM Complete proteome.  
 FT SITE 81 89 "HIGH" REGION.  
 FT SITE 258 262 "KMSKS" REGION.  
 SQ SEQUENCE 375 AA; 43178 MW; 3DBF85DA680F116 CRC64;  
 Query Match 32.7%; Score 735; DB 1; Length 375;  
 Best Local Similarity 41.6%; Pred. No. 2.5e-53;  
 Matches 162; Conservative 79; Mismatches 130; Indels 18; Gaps 11;  
 QY 34 EEDF-VDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFPS 92  
 DB 2 EEFVVTPEVEAGR-----VDYEKLKHFGAKPLTKDEVALLEKAGE--VHPLIRRGFVA 56  
 QY 93 HEDNQVLDAYENKKPFYLYTGRGSSSEAMHGHLPFTFKWLDVFNVLVQMTDDE 152  
 DB 57 HEDFPIKMHQEBGRWALYTGRRGSG--PVHGHVWPVLLKWFSDKFELEVFQITDDE 115  
 QY 153 K-YLWKD-LTLDQAYGDAVENAKDIACGFDINKTFIPSDLDYMGSSGFFYKVVVKIQKAV 211  
 DB 116 KYDDPEKMLEATWMAYNALDVALGSPERHLIIDTKI---KPIYPIAVNVAKTU 172  
 QY 212 TENQVKGIFGFTSDDCIGKISFPALQAPSPNSFPQIFRRTDIQCLIPALIDDPYR 271  
 DB 173 TNNVYATGFGFTSDTNGILIFYPSLIQAVAFPT--ELRREATPV--LIPALIDDPYR 228  
 QY 272 MTRDVAPRIGYKPKALHSTFPALQAGTQKMSADPNSSIFLNDTAKQITKTKNKAHPS 331  
 DB 229 LARDIADALGYKPKPSTLYSKFTMALTG--ESKMSASNPDSALYTLTDEKTVRRKV--MNAFT 286



```

QY 332 GGRDTEIEHROFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTSGLMGLGELKALILEV 391
DB 287 GGRPTAEBOEKRTYGNEVCPVHYHMLFPDDPSVAKIRQDCSKGALCGECKLKHKEKI 346
QY 392 OPLIAHQARKEVTDIEVKEFMTPRKLS 420
DB 347 TFLKEHRRERREKARKG-VDEYRLSVKLS 374

RESULT 12
SYN_HALN1
ID SYN_HALN1 STANDARD; PRT; 380 AA.
AC 09HN66;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE TrpCophany1-cRNA synthetase (EC 6.1.1.2) (TrpCophan--cRNA ligase)
DE (TrpRS).
GN TRPS OR TRPS2 OR VNGS232G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddock D.G., Jablonki P.E., Krebs M.P., Angerling C.M., Dale H.,
RA Isebnager T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam W., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Jiang P., Riley M., Hood L., DasSarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + cRNA(Trp) = AMP +
-1- diphosphate + L-tryptophanyl-cRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AE05109; AAG20355.1; -
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; cRNA-synt_1b.
CC InterPro; IPR001412; cRNA-synt_1.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PR01039; TRNASYNTHRP.
CC TIGR; TIGR00233; trps; 1.
CC PROSITE; PS00178; AA_TRNA_Ligase_I; 1.
CC Aminoacyl-cRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC -----
CC FT SITE 74 82 "HIGH" REGION.
CC SITE 249 253 "KMSKS" REGION.
CC SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;

Query Match 23.8%; Score 534.5; DB 1; Length 380;
Best Local Similarity 35.4%; Pred. No. 9.6e-37;
Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

```

```

DB 54 GSDVDVDFILTAGS-----IVGVGSG--PMILGHAMVYFARLRDDEGARVYVLSDE 108
QY 153 KILMKDLTLDDAYGDAVE-NAKDIIACGPDINKTIF--SDLDYM-GMSGFFYNNVKI 207
DB 109 KWFMDQTPAET-GDYLANLRDLDAVDFDELTRIVVDTRDADVLPLATFADV-- 164
QY 208 QKHVFNQVKGIFGSTDSDCGKISFPALQAPSPNSFPQIFRRTDIOCLIPALIOD 267
DB 165 -RHATLQNVYG----EPDNGQAFYPAVQTRHLL--LPQLVHG--BHEITLVPIAVDD 213
QY 268 PYFRMTDVAAPRIGYP--KPALLHSTFPALQAGTKASAPNSIFLDTAKQIKTV 325
DB 214 PHVRVSRVAAKARYPVGRKPGALLMQFLPSLAG-PGKSSS-AGVISIRLTDSPDIVREXV 271
QY 326 NKHAFSGRDTIEHROFGNCDVVSFWYLTFFLEDDDKLEQIRKDYTSGLMGLGELK 384
DB 272 RTHAVTGRASVEEHRAGVPAEDVPFOYLSAFEPDPAELARIEREVRAGDILSGELK 331
QY 365 KALIEVLOPLIAHQARKEVTDIEVKEFMTPRKLSFD 422
DB 332 DIAADRITFPLAHHQRRAALGD--VTEALDAFRLTDD 367

RESULT 13
SYN_METJA
ID SYN_METJA STANDARD; PRT; 370 AA.
AC 058810;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TrpCophany1-cRNA synthetase (EC 6.1.1.2) (TrpCophan--cRNA ligase)
DE (TrpRS).
GN TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayson R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weisman J.F., Fuhrmann J.L., Nguyen D.,
RA Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii."
RA Science 273:1058-1073(1996).
RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + cRNA(Trp) = AMP +
-1- diphosphate + L-tryptophanyl-cRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC TIGR; MJ1415; -
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; cRNA-synt_1b.
CC InterPro; IPR001412; cRNA-synt_1.
CC Pfam; PF00579; cRNA-synt_1b; 1.
CC PRINTS; PR01039; TRNASYNTHRP.

```



```

RT "Microsporidia, amitochoondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ012470; CAA10034.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 89 98 "HIGH" REGION.
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15744 MW; 60E2935B7E1E344F CRC64;

Query Match 17.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 4,4e-25;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 35 EDFVDPMVTQTSASAK-----GIDYDKLIVRPSSSKIDKELINRIERATGQRPPIFLRGIF 90
DB 3 EGRITPMDEYVUSTDEVPALIDYDKIINRGCEKFNQALADREKLSGKPAHIFFRGIY 62
QY 91 FSHRDNNQVLDAYENKPPVLYTGRGSSSEAMVGHIPFIPTKWLQDVFNVPVLVIQMTD 150
DB 63 FAHRDENLILDEIANNRPFLYLYGRGSSSKTMHIGHTIPFLCKVMQDAFKIRLVITQITD 122
QY 151 DEKYLKQDLTLD 162
DB 123 DEKFLKSMWLE 134

```

Search completed: July 10, 2003, 12:30:34  
 Job time : 10.2509 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:27:20 ; Search time 28.7806 Seconds

(without alignments)  
3035.516 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Sequence: 1 MSYKAAAGSDYKADCEPQNP.....VTDEIVKEMTPRKLSFDRQ 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPTREMBL\_21:\*

- 1: sp archaea:\*
- 2: sp bacteria:\*
- 3: sp fungi:\*
- 4: sp human:\*
- 5: sp invertebrate:\*
- 6: sp mammal:\*
- 7: sp mhc:\*
- 8: sp organelle:\*
- 9: sp phage:\*
- 10: sp plant:\*
- 11: sp rodent:\*
- 12: sp virus:\*
- 13: sp vertebrate:\*
- 14: sp unclassified:\*
- 15: sp rvirus:\*
- 16: sp bacteriophage:\*
- 17: sp archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match length	ID	Description
1	2051	91.3	475	11 Q9DC65	Q9dc65 mus musculus
2	2047	91.1	481	11 Q9J358	Q9j358 mus musculus
3	1629	72.5	329	11 Q9D8R9	Q9d8r9 mus musculus
4	1537	68.4	305	11 Q70184	Q70184 cavia porce
5	1382	61.5	430	5 Q9U4Y1	Q9u4y1 drosophila
6	1379.5	61.4	420	5 Q9U4Y0	Q9u4y0 drosophila
7	1378	61.4	430	5 Q9VHG2	Q9vHG2 drosophila
8	1301	57.9	402	10 Q9SR15	Q9sR15 arabidopsis
9	1027	45.7	324	5 Q9U1R2	Q9u1r2 caenorhabdi
10	942	41.9	385	5 Q8SOY5	Q8sOy5 encephalito
11	826.5	36.8	385	17 Q8U453	Q8u453 pyrococcus
12	800.5	35.6	490	5 Q9U1F5	Q9u1f5 leishmania
13	682	30.4	136	6 Q9S295	Q9s295 sus scrofa
14	388.5	17.3	136	6 Q9T588	Q9t588 bos taurus
15	329.5	14.7	437	17 Q8TUA1	Q8tUA1 methanosarc

17	310	13.8	111	5 Q95YL8	Q95yL8 encephalito
18	294	13.1	107	4 Q9U533	Q9u533 trypanosoma
19	286	12.7	109	4 Q9UD15	Q9ud15 homo sapien
20	274.5	12.2	513	17 Q9HN83	Q9hN83 halobacteri
21	191	8.5	324	17 Q8TX22	Q8tX22 methanopyru
22	188	8.4	364	17 Q9YA64	Q9yA64 aeropyrum p
23	181.5	8.1	331	17 Q979Z1	Q979z1 thermoplasm
24	172	7.7	385	10 Q8S9J2	Q8s9J2 arabidopsis
25	171.5	7.6	351	10 P93018	P93018 arabidopsis
26	167	7.4	351	16 Q9RV06	Q9rv06 deinococcus
27	164.5	7.3	895	10 Q9SGN2	Q9sgN2 arabidopsis
28	158.5	7.1	375	17 Q9V027	Q9v027 pyrococcus
29	156	6.9	372	17 Q8ZT77	Q8zT77 pyrobaculum
30	154.5	6.9	102	11 Q07119	Q07119 halobacteri
31	153.5	6.8	317	17 Q8TS11	Q8tS11 methanosarc
32	152.5	6.8	682	5 Q9N9B8	Q9n9B8 leishmania
33	149.5	6.7	327	17 Q9HN62	Q9hN62 halobacteri
34	149.5	6.7	408	10 P93363	P93363 nicotiana t
35	148	6.6	316	17 Q8ZM77	Q8zM77 pyrobaculum
36	146	6.5	375	17 Q8S739	Q8s739 pyrococcus
37	145	6.5	332	17 Q9HK73	Q9hK73 thermoplasm
38	145	6.5	337	16 Q8Z0T0	Q8z0T0 salmoneila
39	142.5	6.3	375	17 Q8U2H3	Q8u2H3 pyrococcus
40	142.5	6.3	528	11 Q91WQ3	Q91wQ3 mus musculus
41	140	6.2	337	16 Q8ZK00	Q8zK00 salmoneila
42	139	6.2	294	17 Q96VY3	Q96vY3 sulfolobus
43	135	6.0	365	16 Q9KNV7	Q9kNV7 vibrio chol
44	134.5	6.0	525	5 Q9VAV60	Q9vAV60 drosophila
45	130.5	5.8	419	16 Q92BB1	Q92bB1 listeria in

## ALIGNMENTS

RESULT 1	ID	Q9DC65	PRELIMINARY;	PRT;	475 AA.
AC	Q9DC65	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Adult male lung CDNA, RIKEN full-length enriched library, clone:1200002C07, full insert sequence.				
GN	WARS				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=LUNG;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Guetincich S., Hill D., Holtman M., Hume D.A., Kamita M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Sorch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,				
RA	Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse CDNA collection.";				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK004541; BAB23357.1; -				

DR MGD; MGI:104630; Mars.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR001412; tRNA-synt\_1.  
DR InterPro; IPR002306; Tyr tRNA-synt\_1b.  
DR InterPro; IPR000738; WHEP-TRS.  
DR Pfam; PF00579; tRNA-synt\_1d; 1.  
DR Pfam; PF00458; WHEP-TRS; 1.  
DR PRINTS; PR01039; TRNASYNTHTRP.  
DR TIGRFAMs; TIGR00233; tRPS; 1.  
DR PROSITE; PS00178; AA tRNA LIGASE\_I; 1.  
DR PROSITE; PS00762; WHEP-TRS; 1.  
SQ SEQUENCE 475 AA; 53641 MW; C3A67FE85521DE4C CRC64;

	Query Match	Similarity	91.3%	Score	2051	DB	111	Length	475;	
	Best Local	Similarity	90.3%	Pred. No.	4,9e-166;					
	Matches	383;	Conservative	21;	Mismatches	20;	Indels	0;	Gaps	0
QY	1	MSYKAAAGEDYKADCPGPNPAPFTHSHGDPDAIEADPFVDPMVTQVTSASAKGIDYDKLYARE	60							
Db	52	MSYKAAAMEEYKAGCPGPNPAPTRAGRCDSDATASADPFVDPMVTQTSASAKGIDYDKLYAVF	111							
QY	61	GSSKIDKELINRIERATGQRPPIHFLRRGIFFSHRMNNQVLDAENKCKFFYLTTGSGPSS	120							
Db	112	GSSKIDKELINRIERATGQRPPIHFLRRGIFFSHRMNNQVLDAENKCKFFYLTTGSGPSS	171							
QY	121	AMHVGHLPIFIPTKMLQDVFNPLVYIQTDEKYLWKDLTLDOAYGDAVENAKDIIACGF	180							
Db	172	AMHLGHLPLFIPTKMLQDVFNPLVYIOWMSDEKYLWKDLTLDOAASYVENAKDIIACGF	231							
QY	181	DINKTFISDSDIYMGSSGFEYKNVVKIQGHVFNQVKGIFGTFSDDCIGKISFPALQAA	240							
Db	232	DINKTFISDSDIYMGSSGFEYKNVVKIQGHVFNQVKGIFGTFSDDCIGKISFPALQAA	291							
QY	241	SFSNSFPOIFRDRDIOCLIPCAIDODDYPFRMTDVPARIGYPRKALISHTFEPALQAO	300							
Db	292	SFSNSFPIRIFRDRDIOCLIPCAIDODDYPFRMTDVPARIGYPRKALISHTFEPALQAO	351							
QY	301	TKMSASDNNSSIFLVDTPAKOIKTKYKNKAFFSGGRDTIEHHRPGNCQDVVSFWLTFPL	360							
Db	352	TKMSASDNNSSIFLVDTPAKOIKSKKNKAFFSGGRDTIEHHRPGNCQDVVSFWLTFPL	411							
QY	361	EDDDKLEQIRKDYTSGAMLTGELKALILEVLOPLIAEHQARRKEVTDIEVKEFWTPRKLS	420							
Db	412	EDDDKLEQIRKDYTSGAMLTGELKKTLLIDVLOPLIAEHQARRKAVTBEVKEFWTPRKLS	471							
QY	421	EDFO 424								
Db	472	FHFQ 475								

RESULT 2			
099J58			
ID	099J58	PRELIMINARY;	PRT: 481 AA.
AC	099J58:		
DT	01-JUN-2001 (TREMblrel. 17, Created)		
DT	01-JUN-2001 (TREMblrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMblrel. 21, Last annotation update)		
DE	tryptophanyl-tRNA synthetase.		
GN	WARS.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=MAMMARY TUMOR. WAP-TAG MODEL. 5 MONTHS OLD;		
RA	Strausberg R.;		
RL	Submitted (FE8-2001) to the EMBL/Genbank/DBJ databases.		
DR	EMBL: BC003450; AA03450.1; -.		
DR	MGI: MGI:104630; WARS.		
DR	InterPro: IPR002305; tRNA-synt_1b.		
DR	InterPro: IPR001412; tRNA-synt_1.		
DR	InterPro: IPR002306; ttp-tRNA-synt_1b.		

DR	IleRpoPro; IPR000738;	WHEP-TRS.
DR	Pfam; PF00579;	CNA-synt_1b; 1.
DR	Pfam; PF00458;	WHEP-TRS; 1.
DR	PRINTS; PRO1039;	TRNASYNTHRP.
DR	TIGRFAMs; TIGR00233;	tTPS; 1.
DR	PROSITE; PS00178;	AA tRNA LIGASE_I; 1.
DR	PROSITE; PS00762;	WHEP TRS; 1.
KW	Aminacyl-tRNA synthetase.	
SO	SEQUENCE	481 AA; 54325 MW; A754E1DD5F58E2EF3 CnC64;
Query Match	91.1%;	Score 2047; DB 11; Length 481;
Best Local Similarity	90.1%;	Pred. No. 1.le-165;
Matches 382;	Conservative 22;	Mismatches 20; Indels 0; Gaps 0;

Qy	1	MSYFAAGGEDYKADCCPGENPAFTSNHGDAIEAEEDFVDPMFTVOYSSAKGIDYDKLIVRF	60
Db	52	VSYKAAAGGEYKACCPGENPAGENCSDATKASEDFVDPMFTVTRTSAAKGIDYDYLIVQF	111
Qy	61	GSSTKIDELNIRIRATGQRPHHFLRGIFFSHSDMNQVLDAYENKKEFYLYTGSGPSE	120
Db	112	GSSKIDELNIRIRATGQRPHRLRNGIIFPSHSDMNQILDAYENKKEFYLYTGSGPSE	171
Qy	121	AMHGHLPFIPTKMLODVFNVPVLYIOWTDEDEKLMKDLTDQAVGDAVENAKDIACGF	180
Db	172	AMHGHLPVPIFTKMLODVFNVPVLYIQMSDEDEKLMKDLTLEQAVSYVENAKDIACGF	231
Qy	181	DINKTFIFSDLDYMGSSGFYKNVVKIOKHTVFNQVKIGFPTSDICIGKISFPATQAP	240
Db	232	DINKTFIFSDLEVMGSGPGEFYRNVVKIOKHTVFNQVKIGFPTSDICIGKISFPVQAPAP	291
Qy	241	SFSNSFOIPBDRDIOCLPCALIDOPPYPMFTDVAAPRISGPPALHSHFTFPALQAO	300
Db	292	SFSNSFPKIFEDRDIQCLPCALIDODPYPMFTDVAAPRISHPPALHSHFTFPALQAO	351
Qy	301	TKMASASPNSIIFLTDRAKQIKTIVNKHAFSGSGDITIEHROFGGNCVDVDSFWMLTFEFL	360
Db	352	TKMASADPNSSIIFLTDRAKQIKSVNKHAFSGSGDITVEBHQFGGNCVDVDSFWMLTFEFL	411
Qy	361	EDDKLEQIRKDYTSYGAMLTGELKKALILEVLOPLIAEHQARRKEYTDEIVYEFMTPRKLS	420
Db	412	EDDRLEQIRKDYTSYGAMLTGELKKTLIDVLOPLIAEHQARRKEYTDEIVYEFMTPRKLS	471
Qy	421	EDFQ 424	
Db	472	FHFQ 475	

RESULT 3			
09DBR9			
ID	09DBR9	PRELIMINARY;	PRT: 329 AA.
AC	09DBR9		
DT	01-JUN-2001 (TREMBLREL. 17, Created)		
DT	01-JUN-2001 (TREMBLREL. 17, Last sequence update)		
DT	01-MAR-2002 (TREMBLREL. 20, Last annotation update)		
DE	TRYPtophanyl-tRNA synthetase.		
GN	WARS.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Aikawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,		
RA	Aikawa K., Iawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,		
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido T., Peeble G., Quackenbush J.,		
RA	Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Ronaldo M.P.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombarts P.,  
 RA Norioka P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilmink L.,  
 RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK007754; BAB25235.1; -  
 DR MGI:104630; Mavs.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_1b.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA TRNA LIGASE I; 1.  
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244C4A2 CRC64;

Query Match 72.5%; Score 1629; DB 11; Length 329;  
 Best Local Similarity 92.1%; Pred. No. 2,4e-130;  
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 96 MNQVLDAYENKPKFYLYTGRGSSSEAMHGHLPFTKWLQDVNVPVLTQMTDEKYL 155  
 DB 1 MNQVLDAYENKPKFYLYTGRGSSSEAMHGHLPFTKWLQDVNVPVLTQMTDEKYL 60  
 QY 156 MKDLTLDQAYGVDAVENAKDIIACGPDINKTFIFSDIDYMGSSGFKYKVVKIOKHTVFNQ 215  
 DB 61 MKDLTLDQAYGVDAVENAKDIIACGPDINKTFIFSDIDYMGSSGFKYKVVKIOKHTVFNQ 120  
 QY 216 VKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAIDQDPYFMRTD 275  
 DB 121 VKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCAIDQDPYFMRTD 180  
 QY 276 VAPRIGYPPALHSTFFPALQAGOTKMSADPNSSIFLTDTKQIKTKVKNKAFSGGRD 335  
 DB 181 VAPRIGYPPALHSTFFPALQAGOTKMSADPNSSIFLTDTKQIKTKVKNKAFSGGRD 240  
 QY 336 TTEEHQFGGNCVDVSMYLTFFLEDDDLQEIIRKDYTSGLMTGELKKALIEVLQPLI 395  
 DB 241 TTEEHQFGGNCVDVSMYLTFFLEDDDLQEIIRKDYTSGLMTGELKKALIEVLQPLI 300  
 QY 396 AEHQARKKEVTDEIVKEFMTPRLSFDFQ 424  
 DB 301 AEHQARKKEVTDEIVKEFMTPRLSFDFQ 329

RESULT 4  
 ID 070184 PRELIMINARY; PRT; 305 AA.  
 AC 070184;  
 DT 01-AUG-1998 (TEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)  
 DE Trypophan-tRNA synthetase (Fragment).  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.  
 NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HARTLEY; TISSUE=SPLEEN;  
 RA Yang D., Goto R., Watanabe N., Kobayashi Y.;  
 RT "Identification and Cloning of Genes Whose Expressions are Elevated  
 RT during DNCB-induced Guinea Pig Skin Delayed-type Hypersensitivity  
 RT Reaction";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB012222; BAA25288.1; -  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR002306; tRNA-synt\_1b.

DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 KW Aminoacyl-tRNA synthetase.  
 FT NON-TER  
 SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 68.4%; Score 1537; DB 11; Length 305;  
 Best Local Similarity 94.4%; Pred. No. 1,4e-122;  
 Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 120 EAMHVGHLIPFTKWLQDVNVPVLTQMTDEKYLMDLTLDQAYGVDAVENAKDIIACG 179  
 DB 1 EAMHVGHLIPFTKWLQDVNVPVLTQMTDEKYLMDLTLDQAYGVDAVENAKDIIACG 60  
 QY 180 FQINKTFIFSDIDYMGSSGFKYKVVKIOKHTVFNQKIFFTSDICIGKISFPALQAA 239  
 DB 61 FQINKTFIFSDIDYMGSSGFKYKVVKIOKHTVFNQKIFFTSDICIGKISFPALQAA 120  
 QY 240 PPSNSFPQIFRDRDIOCLIPCAIDQDPYFMRTDVAPRIGYPPALHSTFFPALQGA 299  
 DB 121 PPSNSFPQIFRDRDIOCLIPCAIDQDPYFMRTDVAPRIGYPPALHSTFFPALQGA 180  
 QY 300 QTKMSADPNSSIFLTDTKQIKTKVKNKAFSGGRDITTEEHQFGGNCVDVSMYLTFF 359  
 DB 181 QTKMSADPNSSIFLTDTKQIKTKVKNKAFSGGRDITTEEHQFGGNCVDVSMYLTFF 240  
 QY 360 LEDDDLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKEVTDEIVKEFMTPRKL 419  
 DB 241 LEDDDLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARKKEVTDEIVKEFMTPRKL 300  
 QY 420 SFDQ 424  
 DB 301 SFDQ 305

RESULT 5  
 ID 0904Y1 PRELIMINARY; PRT; 430 AA.  
 AC 0904Y1;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE Trypophanyl-tRNA synthetase (AT21437p).  
 GN AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99250164; PubMed=10233165;  
 RA Seshiah P., Andrew D.J.;  
 RT "WRS-85D: A trypophanyl-tRNA synthetase expressed to high levels in  
 RT the developing Drosophila salivary gland.";  
 RL Mol. Biol. Cell 10:1595-1608(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dreesen D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guartin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunco J., Paclob J., Paragas V., Park S.,  
 RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF125156; AAF20166.1; -  
 DR EMBL: AY075249; AAL68116.1; -  
 DR FLYBase: FBgn0010803; Aats-trp.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_1b.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.

DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;  
 Query Match 61.5%; Score 1382; DB 5; Length 430;  
 Best Local Similarity 61.6%; Pred. No. 3.6e-109;  
 Matches 263; Conservative 63; Mismatches 85; Indels 16; Gaps 3;

QY 10 DYKADCEPQNPAPTSNHPDA-----TEAE-----EDFVDEPTVOTSSAKGIDYD 54  
 DB 3 DTKETVEGEVETGTDAAOEGATAPTEDVDDPNNVASSNDAGVDYD 62  
 QY 55 KLIVREGSSKIDKELINIRIRATGQRPHHFLRRGIFPSHRDMMOVLDAENKKPFYLTYTG 114  
 DB 63 KLIRFSSSKIDKELINIRIRATGQRPHHFLRRGIFPSHRDMMOVLDAENKKPFYLTYTG 122  
 QY 115 RGSSEAMHVGHLIPFIETKMLQDVFNVPVLTQMTDEKYLMDLTDQAVGDAVENAKD 174  
 DB 123 RGSSEAMHVGHLIPFIETKMLQDVFNVPVLTQMTDEKYLMDLTDQAVGDAVENAKD 182  
 QY 175 IIAAGPDKNTFFISDLDYMGSSGFKYKVVKIOKAVTNQVKGIFGFTSDGKISFP 234  
 DB 183 IVAIGFVNKTFIFNNLEFVKGCPAMYNITRIQCVTFNQVKGIFGFGSDIIGKIGFP 242  
 QY 235 AIAAASFNSPQIPFRDRTDIOCLIPCAIDDDPYFRMTDVAIRIGYPRKALLHSTFP 294  
 DB 243 AIAAASFNSPQIPFRDRTDIOCLIPCAIDDDPYFRMTDVAIRIGYPRKALLHSTFP 301  
 QY 295 AIOGAQTKMSASDPNSSIFLTDTKAKOIKTKVKNKIASGSGDITIEHRQFGNCDVVSFM 354  
 DB 302 AIOGAQTKMSASDPNSSIFLTDTKAKOIKTKVKNKIASGSGDITIEHRQFGNCDVVSFM 361  
 QY 355 YLFFLEDDEKLEIRKDYTSGLMTSELKKALIEVLOPLAEHQARKKEVTDEIVAEFM 414  
 DB 362 LKFFLEDDEKLEIRKDYTSGLMTSELKKALIEVLOPLAEHQARKKEVTDEIVAEFM 421  
 QY 415 TPKRLSF 421  
 DB 422 ELAPLKR 428

RESULT 6  
 Q9U4Y0 PRELIMINARY; PRT; 420 AA.  
 AC Q9U4Y0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (Fragment).  
 GN AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RX MEDLINE=99250164; PubMed=10233165;  
 RA Seshiah P., Andrew D.J.;  
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in  
 the developing Drosophila salivary gland.";  
 RL Mol. Biol. Cell 10:1595-1608(1999).  
 DR EMBL; AF125157; AAF20167.1; .  
 DR FlyBase; FBgn0010803; Aats-trp.  
 DR InterPro; IPR002305; CRNA-synt\_1b.  
 DR InterPro; IPR004142; CRNA-synt\_1.  
 DR InterPro; IPR002306; Trp CRNA-synt\_1b.  
 DR Pfam; PF00579; CRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASNTTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.

FT NON TER. 1 1  
 SQ SEQUENCE 420 AA; 46975 MW; 4AF7088AF426AB6 CRC64;  
 Query Match 61.4%; Score 1379.5; DB 5; Length 420;  
 Best Local Similarity 62.7%; Pred. No. 5.6e-109;  
 Matches 257; Conservative 62; Mismatches 90; Indels 1; Gaps 1;

QY 12 KADCPGNPAPTSNHPDTEAEDEDVDPMTVOTSSAKGIDYDKLIVFGSSKIDKELIN 71  
 DB 10 KPDAEVEETGTDAAOEGATAPTEDVDDPNNVASSNDAGVDYDKLIRFSSSKIDKELIN 69  
 QY 72 RIRATGQRPHHFLRRGIFPSHRDMMOVLDAENKKPFYLTYTGRSSSEAMHVGHLIPFI 131  
 DB 70 RFIKIGKAPHHIRRGIMFSSHRDLHTLTLRQKGFYLYTGRSSSGLHVGHLVPI 129  
 QY 132 FTKMLQDVFNVPVLTQMTDEKYLMDLTDQAVGDAVENAKDIIACGPDKNTFFISDL 191  
 DB 130 MTKMLQETFPVPLVLTQMTDEKYLMDLTDQAVGDAVENAKDIIACGPDKNTFFISDL 189  
 QY 192 DYMGSSGFKYKVVKIOKAVTNQVKGIFGFTSDGKISFPALGAAPSNSPQIPFR 251  
 DB 190 EFKGKCPAMYNITRIQCVTFNQVKGIFGFGSDIIGKIGFPAAQAPASISTFFIFG 249  
 QY 252 DRDIOCLIPCAIDDDPYFRMTDVAIRIGYPRKALLHSTFPALGAQTKMSASDPNSS 311  
 DB 250 NR-KVHCLIPCAIDDDPYFRMTDVAIRIGYPRKALLHSTFPALGAQTKMSASDPNSSA 308  
 QY 312 IFETIDPAKQIKTKVKNKIASGSGDITIEHRQFGNCDVVSFMYLFFLEDDEKLEIRK 371  
 DB 309 VYLLTDPEKQIKTKVKNKIASGSGDITIEHRQFGNCDVVSFMYLFFLEDDEKLEIRK 368  
 QY 372 DYSGAMLTGELKKALIEVLOPLAEHQARKKEVTDEIVAEFMTPKRLSF 421  
 DB 369 AYSKGMTLGEIKLAVETLTPIVEHQARKKITBEVLDKYEAPLKR 418

RESULT 7  
 Q9VHG2 PRELIMINARY; PRT; 430 AA.  
 AC Q9VHG2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Aats-trp protein.  
 GN AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang C., Chen L.X.,  
 Brandon R.C., Rogers J.H.C., Blazer R.G., Champ M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Berooz P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 Burdus K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evansgelist C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,



RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mantei B., McIntosh T.C., Meled M.P., Mepheson D.,  
 RA Merkulov G., Mishina N.V., Mobary C., Morris J., Moshel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stepleton M., Strong R., Sun B.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan G., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003682; AAG22136.1; -;  
 DR FLYBase; FBgn0010803; Atcs-tp.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; tTP\_cRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 DR SEQUENCE 430 AA; 47971 MW; 2938EECC69E979F CRC64;

Query Match 61.4%; Score 1378; DB 5; Length 430;  
 Best Local Similarity 61.4%; Pred. No. 7,8e-109;  
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

10 DYKADCPGNPATSHGPD-----TEAE-----EDFVDPWVQTSSAKIDYD 54  
 DB 3 DKETVVEGEALTLNGKPDABEVEGTGTDQAQEGATAPTEDEVDPMNVAASNDAGVDYD 62  
 QY KLIIVRGSSKIDELINIRERATGQRPHFLRGIFPSHNDQNVLDAYENKKEPFLYNG 114  
 DB 63 KLIIVRGSSKIDELINIRERATGQRPHFLRGIFPSHNDQNVLDAYENKKEPFLYNG 122  
 QY 115 RGGSSAMVGHVLIPIFTKMLQDVFNPLVLIQMTDEKYLKDLTLDQAYGAVENAKD 174  
 DB 123 RGGSSAMVGHVLIPIFTKMLQDVFNPLVLIQMTDEKYLKDLTLDQAYGAVENAKD 182  
 QY 175 IIAAGDINKTIFSDLDYWGSSGFYKAVVVKIQKHTENVQGIIFGTTSDICIGISFP 234  
 DB 183 IIAAGDINKTIFSDLDYWGSSGFYKAVVVKIQKHTENVQGIIFGTTSDICIGISFP 242  
 QY 235 AIOAASFNSPFOIRDRPTDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 294  
 DB 243 AIOAASFNSPFOIRDRPTDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 301  
 QY 295 AIOAASFNSPFOIRDRPTDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 354  
 DB 302 AIOAASFNSPFOIRDRPTDLOCLPCALDODPYFRMRDVAIRIGYKPKALHSTFP 361  
 QY 355 YLTFEEDDKLQIKRKYTGSMALTEGLKALILEVLQPLIAHQARKKVEVDIYKEM 414  
 DB 362 YLTFEEDDKLQIKRKYTGSMALTEGLKALILEVLQPLIAHQARKKVEVDIYKEM 421  
 QY 415 YLTFEEDDKLQIKRKYTGSMALTEGLKALILEVLQPLIAHQARKKVEVDIYKEM 421  
 DB 422 YLTFEEDDKLQIKRKYTGSMALTEGLKALILEVLQPLIAHQARKKVEVDIYKEM 428

RESULT 8  
 Q9SR15 PRELIMINARY; PRT; 402 AA.  
 AC Q9SR15, 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Putative tryptophanyl-tRNA synthetase.  
 GN F7018.7 OR AF3604600.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsie.  
 OX NCBI\_TaxId=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,  
 RA Roming C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
 RA Bowman C.L., White O., Niernan W.C., Frazer C.M.,  
 RT "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tortum M., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai U., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.D., Sakurai T., Satou M.,  
 RA Seki M., Shimizu P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis full length cDNA clones.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC011437; AAP04890.1; -;  
 DR EMBL; AY080709; AAL85027.1; -;  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; tTP\_cRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; ttps; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 DR Aminoacyl-tRNA synthetase.  
 DR SEQUENCE 402 AA; 45754 MW; B9BEA75E5D6CD15 CRC64;

Query Match 57.9%; Score 1301; DB 10; Length 402;  
 Best Local Similarity 60.3%; Pred. No. 2.5e-102;  
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

29 DATAE--EDFDPWVQTSSAKIDYDYLIRFGSSKIDKLINIRERATGQRPHFLR 86  
 DB 7 DEREAASSQVNVNPEVSAKDGKIDYDKLIDFGQRLDESILDRVORLTSRPHVFLR 66  
 QY 87 RGIIFSHRDMNOVLDAVENKKEPFLYLTGSPSEAMVGHVLIPIFTKMLQDVFNPLVI 146  
 DB 67 RGIIFSHRDMNOVLDAVENKKEPFLYLTGSPSEAMVGHVLIPIFTKMLQDVFNPLVI 126  
 QY 147 QMTDEKYLKDLTLDQAYGAVENAKDIIAAGDINKTIFSDLDYWGSSGFYKAVVVK 206  
 DB 127 QMTDEKYLKDLTLDQAYGAVENAKDIIAAGDINKTIFSDLDYWGSSGFYKAVVVK 184  
 QY 207 ICKHTENVQKIFGTTSDICIGISFPALQABSPNSPFOIRDRPTDLOCLPCALD 266  
 DB 185 VIKCVTLNKAAGIFGSGEDPTAKLSFPVQVNPSPSPSPHLPFGKDLRLCLIPALDQ 244  
 QY 267 DVFPMTRDVAIRIGYKPKALHSTFPALQAGTKMSASDNNISIFLDTAKQIKTKVN 326  
 DB 245 DVFPMTRDVAIRIGYKPKALHSTFPALQAGTKMSASDNNISIFLDTAKQIKTKVN 304  
 QY 327 KIAFSGGRTDTEHRQFGNCDVDSFWYLTFLFEDDDKLEQIRQDYTGAMLTGELKXA 386  
 DB 305 RYAFSGGRTDTEHRQFGNCDVDSFWYLTFLFEDDDKLEQIRQDYTGAMLTGELKXA 364  
 QY 387 LIEVLQPLIAHQARKKVEVDIYKEMTPRLTSPDQ 424  
 DB 365 LIEVLQPLIAHQARKKVEVDIYKEMTPRLTSPDQ 402

RESULT 9

```

Q9U1R2
ID Q9U1R2 PRELIMINARY; PRT; 324 AA.
AC Q9U1R2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Y80D3A.1 protein.
GN Y80D3A.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
NCBI_TaxID=6239;
OX [1]
RN SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating Biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL132853; CAB60439.1; -.
DR InterPro; IPR002305; CRNA-synt_1b.
DR InterPro; IPR002306; TTP_CRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASyntTRP.
SQ SEQUENCE 324 AA; 36289 MW; 6E687E6E420EC12 CRC64;

Query Match 45.7%; Score 1027; DB 5; Length 324;
Best Local Similarity 64.9%; Pred. No. 3.8e-79;
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

QY 27 GPDATAEEDPVPFWTVOTSSAKIDYDKLIVRGSSKIDKELINRIERATGQRPHFLR 86
DB 23 GGGQVEDDEEDRVETWETTKATGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLR 82
QY 87 RGFIFSRDMNQVIDAENKKPFYLYTGRGSSAMVGHILPIFKMLQDVNVNVLVI 146
DB 83 RGFIFSRDMNQVIDAENKKPFYLYTGRGSSAMVGHILPIFKMLQDVNVNVLVI 142
QY 147 QMTDDEKYLWKLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDYMGSSGFYKVVVK 206
DB 143 QMTDDEKYLWKLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDYMGSSGFYKVVVK 200
QY 207 IQKXVTFNQVKGIFGFTDSDICGISFPALQAPSFNSFPQIFRRTDIQCLIPCAIDQ 266
DB 201 IMKVVNTNQARAIFFGFTPEDCIGAKAPPAVEAACPFASSFPQIFGKENDIPCLIPCAIDQ 260
QY 267 DPFYFRMTRDVAAPRIGYKAPALHSTFPFALQAGQTKASDPNSIFLTTAKQIKTKV 325
DB 261 DPFYFRMTRDVAAPRIGYKAPALHSTFPFALQAGQTKASDPNSIFLTTAKQIKTKV 319

RESULT 10
Q8S0Y5 PRELIMINARY; PRT; 385 AA.
AC Q8S0Y5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tyrocinophanyl tRNA synthetase.
GN EC011_0530.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
NCBI_TaxID=6035;
OX [1]
RN SEQUENCE FROM N.A.
RA STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prenster G., Barbe V., Peyrethallade E., Broctier P., Wincker P.,
RA Delbec F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vavres C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590450; CAD25963.1; -.
SQ SEQUENCE 385 AA; 44188 MW; 2865C3AC95FCC859 CRC64;

Query Match 41.9%; Score 942; DB 5; Length 385;
Best Local Similarity 46.1%; Pred. No. 8.3e-72;
Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

QY 35 EDPVDPWTVOTSSAK----GIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGIF 90
DB 3 EQRTFMWDEVEVSTDEVPVADYDKLINQGCCKEKNQALADREKLSGKPAHYFFRRGIY 62
QY 91 FSHRDMNQVIDAENKKPFYLYTGRGSSAMVGHILPIFKMLQDVNVNVLVIQMTD 150
DB 63 FAHRDFNLIDELANNRPFLYLYTGRGSSKTMHIGHTIPPLCKYQDAKRLVLIQITD 122
QY 151 DEKYLWKLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDYMGSSGFYKVVVKIQKH 210
DB 123 DEKYLWKLTLDOAYGDAVENAKDIIACGFIDNKTFFISLDYMGSSGFYKVVVKIQKH 179
QY 211 VTFNQVKGIFGFTDSDICGISFPALQAPSFNSFPQIFRRTDIQCLIPCAIDDPYF 270
DB 180 ININEMAIKVGFMSSNIGVGPFAKEIAPCFSSFRFGIK---GAMCLVPAVADDPFF 236
QY 271 RMTRDVAAPRIGYKAPALHSTFPFALQAGQTKASADPNSSIFLTTAKQIKTKVNAF 330
DB 237 RLARDKAKALGEEKKPSIYSLPLDKGVKRRKASADPNSSIFLTTAKQIKTKVNAF 296
QY 331 SGGRTDIEEHROGNCQDVVSFMYLTFLEDDDKLEQIRKQVTSQAMLTGELKKALIEV 390
DB 297 SGGRTDIEEHREKGGIDVDVPEYLYKFLDDDOELKRSKGIKEIITSKEMKCKCVV 356
QY 391 LQPLIAHQARKEVTDEIVEKEMTPRK 418
DB 357 IQGFVSRYGQSRKRVTDLDLRAFIDINK 384

```

```

RESULT 11
Q8U453 PRELIMINARY; PRT; 385 AA.
AC Q8U453;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tyrocinophanyl tRNA synthetase.
GN PF0241.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RN SEQUENCE FROM N.A.
RA STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010149; AL180365.1; -.
KW Aminocacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 385 AA; 45178 MW; 3A7A628958200CCC CRC64;

Query Match 36.8%; Score 826.5; DB 17; Length 385;
Best Local Similarity 45.9%; Pred. No. 5.6e-62;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

```

```

QY 34 BEDF-VDPMWVOTSSAKGIDYDLIVRFGSSKIDKELINRERATQGRPHHFLRGIFPS 92
DB 2 BEEFKTPMEVEI-----IDYNKLIQFGFSPILDTLERTALITSELPFRFRKFFPS 57
QY 93 HEDMNOVLDAYENKPEFVLTGSGPSSSEAMVGHLLPFIFTKMLQDVFNVLQMTDDE 152
DB 58 HEDYDVLDVDEYQKGFVLTGSGPSSG-PMHIGHIIPFATKMLQKFDVNLVLTQTTDE 116
QY 153 KTLMD-LTLDOAYGAVENAKDIIACGPIINTFFISDLDPYMGSSGFYKVVKIQIV 211
DB 117 KFLFKNLFFEDTKVAYONIIDILAVGPDPTFFLQNSEF-----TKLYEAMAPIAKKI 172
QY 212 TENOYVGIFFGFTSDICIGKISFPALQAAFSFNSFPQIFRDRDIOCLIPCAIDODPYR 271
DB 173 NTSMAKAVGFTGFSKIGMIFPAIQAAPIF-----FEK-----RCLIPALDODPYR 222
QY 272 MTRDVAIRIGYKPKALLHSTFPALQAAQOTKMSASDPNSSIFLTDKAKOIKTKVNHAFS 331
DB 223 LORDFAESLIGYVYKTAALHSKFPPLTGLRGKMSASKEPATAVLTDPNPEAGKKIMWFALT 282
QY 332 GGRDTIEHRQRCNGCDVVSFMYLTFLEDDDKLEQIRKY---TSGMLTGLKAL 388
DB 283 GGQPLTKERKRGKNEKCVFEMLEIFPEPDK--KLMEYVACKNGELLCGECKRYII 340
QY 389 EVLQPLIAHQAARKKVTDEIVKEFMTPRKL 420
DB 341 QVQEFLEKHEQKRGK-AETLVKFKYTKGLA 371

```

## RESULT 12

```

ID 09U1F5 PRELIMINARY; PRT; 490 AA.
AC 09U1F5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Tryptophanyl-tRNA synthetase.
GN L1063.04.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,
RA Barrell B.G.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;
RX MEDLINE=98146435; PubMed=9477341;
RX Ivens A.C., Lewis S.M., Bagnerzdeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RA "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL121862; CAB58393.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b.
DR PRINTS; PRO1039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 490 AA; 54904 MW; 172068622C511D3D CRC64;

```

Query Match 35.6%; Score 800.5; DB 5; Length 490;  
 Best Local Similarity 37.7%; Pred. No. 1.3e-59;  
 Matches 174; Conservative 89; Mismatches 133; Indels 65; Gaps 11;

```

QY 19 NPAFTNHPGDAIEADPDFTVOTSSAKGIDYDLIVRFGSSKID-----KELIN 71
DB 33 DPSEPVQHPDDGAGADVITPWWVAAKGPQGINRVLTIFKAERMDGARGHMDVNA 92

```

```

QY 72 RIERAT-----GGRP-----HFLRGCI 89
DB 93 KCRKRTMTTSPSANNAGVATPVLDVAPGADALQAAAPMGOQROQPPALALHHPHRDI 152
QY 90 FFSHDMNOVLDAYE-----NKKPFYLTGSGPSSSEAMVGHLLPFIFTKMLQDVFNVL 144
DB 153 ASHRDLHVALVDIEASVKTGRSVFLYTGROPSAGTMHLGHVLPMLTKYLQDVESLPL 212
QY 145 VQMTDDEKYLKDLTLDOAYGDV--ENAKDIIACGPIINTFFISDLDPYMGSSGFYK 202
DB 213 VQITDDEKFLRDPVFEKAKADELIRSNIKDIIENFPRHTFLFRNHYG--DMWP 269
QY 203 NVVLTOKHTVFNQVGIFFGFTSDICIGKISFPALQAAFSFNSFPQIFRDRDIOCL 260
DB 270 TVLRQDSMTGNAVHTLGITSDVNGKLAFAPTQAAPECFSTAFKRVLRNGDR-PKRCII 328
QY 261 PCAIDODPYFRMTRDVAIRIGYKPKALLHSTFPALQAAQOTKMSAS-DPNSSIFLTDK 319
DB 329 PCAIDODPEFVLTTRAALRLKQLPALHHTKFLPALKGLHGMSSSAEKGVITLHDTDK 388
QY 320 QIKTKVNHAFSGGRDTIEHRQRCNGCDVVSFMYLTFLEDDDKLEQIRKYSGAML 379
DB 389 QVRKRL-RRAFGGCAITLQOMQGTANLELDVAYOYLRFCDPTLFLADVTQRYRGTIN 447
QY 380 TSELKALTE-VLQPLIAHQAARKKVTDEIVKEFMTPRKL 419
DB 448 SEVVDLADLCIRREVLDHMRERRATVDDDVYEFGRIRDI 488

```

## RESULT 13

```

ID 095295 PRELIMINARY; PRT; 136 AA.
AC 095295;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Tryptophanyl-tRNA synthase (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Wintneroe A.K., Fredholm M., Davies W.;
RT Evaluation and characterization of a porcine small intestine cDNA
RT library.
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81267; CAB03585.1; -.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR002306; tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b.
DR PRINTS; PRO1039; TRNASYNTHTRP.
FT NON TER 1 1
FT NON TER 136 136
SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

```

Query Match 30.4%; Score 682; DB 6; Length 136;  
 Best Local Similarity 94.1%; Pred. No. 2.5e-50;  
 Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

QY 234 PALQAAFSFNSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTDF 293
DB 1 PALQAAFSFSSFPQIFRDRDIOCLIPCAIDODPYFRMTRDVAIRIGYKPKALLHSTDF 60
QY 294 PALQAAQOTKMSASDPNSSIFLTDKAKOIKTKVNHAFSGGRDTIEHRQRCNGCDVVSF 353
DB 61 PALQAAQOTKMSAXAPNSSIFLTDKAKOIKTKVNHAFSGGRDTIEHRQRCNGCDVVSF 120
QY 354 MYLTFLEDDDKLEQI 369
DB 121 MYLTFLEDDDKLEQI 136

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:27:20 ; Search time 27.2194 Seconds

(without alignments)  
3035.516 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGDPATAEDFDVDPWTV.....VTDEIYKEMTPRKLSFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1957	92.5	475	11	Q9DC65
2	1957	92.5	481	11	Q99J58
3	1629	77.0	329	11	Q9DBR9
4	1537	72.6	305	11	Q70184
5	1374.5	65.0	420	5	Q9UAY0
6	1374.5	65.0	430	5	Q9UAY1
7	1370.5	64.8	430	5	Q9VHG2
8	1301	61.5	402	10	Q9SR15
9	1027	48.5	324	5	Q9UR12
10	942	44.5	385	5	Q8SOY5
11	826.5	39.1	385	17	Q8U453
12	795.5	37.6	490	5	Q9UUF5
13	682	32.2	136	6	Q9S295
14	453	21.4	374	17	Q8TYF7
15	329.5	15.6	437	17	Q8TUAI
16	310	14.7	111	5	Q95YL8
					Q9GC65 mus muscul
					Q99J58 mus muscul
					Q9DBR9 mus muscul
					Q70184 cavia porc
					Q9UAY0 drosophila
					Q9UAY1 drosophila
					Q9VHG2 drosophila
					Q9SR15 arabidopsi
					Q9UR12 caenorhabdi
					Q8SOY5 encephalito
					Q8U453 pyrococcus
					Q9UUF5 leishmania
					Q9S295 sus scrofa
					Q8TYF7 methanopyr
					Q8TUAI methanospir
					Q95YL8 encephalito

17	296.5	14.0	136	6	Q9TS88	Q9TS88 bos taurus
18	294	13.9	157	5	Q9US33	Q9US33 trypanosoma
19	286	13.5	109	4	Q9UD15	Q9UD15 homo sapien
20	269.5	12.7	513	17	Q9HN83	Q9HN83 halobacteri
21	191	9.0	324	17	Q8TX22	Q8TX22 methanopyru
22	188	8.9	364	17	Q9YA64	Q9YA64 aeropyrum p
23	181.5	8.6	331	17	Q979Z1	Q979Z1 thermoplasm
24	171.5	8.1	351	10	P93018	P93018 arabidopsis
25	171.5	8.1	385	10	Q8S9J2	Q8S9J2 arabidopsis
26	167	7.9	351	16	Q9RVDS	Q9RVDS deinococcus
27	164.5	7.8	895	10	Q9SGN2	Q9SGN2 arabidopsis
28	158.5	7.5	375	17	Q9V027	Q9V027 pyrococcus
29	156	7.4	372	17	Q8ZY77	Q8ZY77 pyrobaculum
30	154.5	7.3	102	1	O07119	O07119 halobacteri
31	153.5	7.2	317	17	Q8TS11	Q8TS11 methanosarc
32	152.5	7.2	682	5	Q9N9B8	Q9N9B8 leishmania
33	149.5	7.1	408	10	Q9HN62	Q9HN62 halobacteri
34	149.5	7.1	408	10	P93363	P93363 nicotiana t
35	148	7.0	316	17	Q8ZW77	Q8ZW77 pyrobaculum
36	146	6.9	375	17	O58739	O58739 pyrococcus
37	145	6.9	332	17	Q9HK73	Q9HK73 thermoplasm
38	145	6.9	337	16	Q8Z0Y0	Q8Z0Y0 salmomella
39	142.5	6.7	375	17	Q8U2H3	Q8U2H3 pyrococcus
40	142.5	6.7	528	11	Q91WQ3	Q91WQ3 mus musculu
41	140	6.6	337	16	Q8ZK00	Q8ZK00 salmomella
42	139	6.6	294	17	Q96YV3	Q96YV3 sulfobius
43	135	6.4	365	16	Q9KRV7	Q9KRV7 vibrio chol
44	134.5	6.4	525	5	Q9VAV0	Q9VAV0 dirosophila
45	130.5	6.2	419	16	Q92BB1	Q92BB1 listeria in

## ALIGNMENTS

RESULT 1

ID Q9DC65 PRELIMINARY; PRT; 475 AA.

AC Q9DC65;

DT 01-JUN-2001 (TEMBLrel. 17, Created)

DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)

DE Adult male lung cDNA, RIKEN full-length enriched library,

DE clone:1200002C07, full insert sequence.

GN WARS.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

ON NCBI\_TaxId=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=LUNG; MEDLINE=21085660; PubMed=11217851;

RX Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T., Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Grotstein S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H., Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL, AK004541, BAB23357.1, -.

DR MED; MGI: 104630; WARS.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; Ttd tRNA-synt\_1b.  
DR InterPro: IPR00738; WHEP-TRS.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR Pfam: PF00458; WHEP-TRS; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
DR TIGRFAMs: TIGR00233; tRPS; 1.  
DR PROSITE: PS00178; AA tRNA LIGASE\_I; 1.  
DR PROSITE: PS00762; WHEP-TRS; 1.  
SQ SEQUENCE 475 AA; 53641 MW; C3467EB85621DB4C CRC64;

Query Match	92.5%	Score 1957	DB 11	Length 475
Best Local Similarity	91.5%	Pred. No. 3.6e-159		
Matches	366	Conservative	20	Mismatches 14; Indels 0; Gaps 0

  

QY	2	NHGPDATEAEEDFVDMPTVQTSSAKGIDYKLIIVFGSSKIDKELINRIERATGQRPHE	61
DB	76	NCSDATASASEDFVDMPTVQTSSAGIDYDKLIIVFGSSKIDKELINRIERATGQRPHE <td>135</td>	135
QY	62	LRRGIFFSHRMNOVLDAVENKKPEYLTGRGSSSEAMHVGILPFITKMLQDVFNPL	121
DB	136	LRRGIFFSHRMNOVLDAVENKKPEYLTGRGSSSEAMHVGILVPIITKMLQDVFNPL	195
QY	122	VIQMTDEKYLMDKLTLLDOAAYGDAVENAKDIIACFDDINTKFIIESDLDYMGMSGFYKVV	181
DB	196	VIQMSDEKYLMDKLTLLDOAYSIVYENAKDIIACGFDDINTKFIIESDLEVMGMSPGFYKVV	255
QY	182	VKIQKHTVFNQVKGIFGFTTSDDCIGKISFPALQAAFPSNSGFPQFRORTDIQCLIPCAI	241
DB	256	VKIQKHVFNQVKGIFGFTTSDDCIGKISFPAAQAPSPNSGPKFRORTDIQCLIPCAI	315
QY	242	DQDQYFRMTROVAPRIGYKPEPALHSTFEPALOGAQOTMSADPSSIFLDTAKOIKTK	301
DB	316	DQDQYFRMTROVAPRIGHPKPALHSTFEPALOGAQOTMSADPSSIFLDTAKOIKSK	375
QY	302	VNKAFAFGSGRGTIEBHROFGANGCDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGELK	361
DB	376	VNKAFAFGSGRGTIEBHROFGGNGCEVDVSFMYLTFFLEDDDRLEQIRKDYTSGAMLTGELK	435
QY	362	KALIEVLOPLIAEHQARRKEVTDLEIVKSPMPKXISPDQ	401
DB	436	KTLIDVLOPLIAEHQARRKAVTEEVKSPMPKXISPDQ	475

RESULT 2	ID	Q99J58	PRELIMINARY;	PRT;	481 AA.
AC	Q99J58;				
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	tryptophanyl-tRNA synthetase.				
GN	WARS.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus				
OX	NCBI_TaxId=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=MAMMARY TUMOR. WAP-TAG MODEL. 5 MONTHS OLD.				
RA	Strausberg R.;				
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: BC003450.1; AA03450.1; -				
DR	MED; MGJ:104630; Waze.				
DR	InterPro; IPR002305; tRNA-synt_1b.				
DR	InterPro; IPR001412; tRNA-synt_1.				
DR	InterPro; IPR002306; Try tRNA-synt_1b.				
DR	InterPro; IPR000738; WHEP-TRS.				
DR	Pfam; PF00579; tRNA-synt_1b; 1.				
DR	Pfam; PF00458; WHEP-TRS_1.				
DR	PRINTS; PRO1039; TRANSYTHTRP				

DR TIGRFAMS: TIGR00233; tps; 1.  
DR PROSITE: PS00178; AA TRNA LIGASE\_I; 1.  
DR PROSITE: PS00762; WHEP TRS; 1.  
KW Aminoacyl-tRNA synthetase  
SEQUENCE 481 AA; 5435 MW; A754E1DDF58E2EF3 CRC64;

	Query Match	Similarity	92.5%;	Score 1957;	DB 11;	Length 481;
	Best Local	Similarity	91.5%;	Pred. No. 3,76-159;		
	Matches	366;	Conservative	20;	Mismatches 14;	Indels 0; Gaps 0;
Qy	2	NHGDATEAEDEDFPMTVMTVOTSSAKGIDYDKLYIRFGSSKIDKELINRIEATGQRPHNF	61			
Db	76	NCSDADKAKASEDFPDPMTVTRTSSAKGIDYDYLIVQFGSSKIDKELINRIEATGQRPHNF	135			
Qy	62	LRGIFGFSHRMNQVLAAYENKKEPFYLYTGRGSPSEAMHVGHLIPFTTKLQDVFNVL	121			
Db	136	LRKGIFFSHRDMQILDAAYENKKEPFYLYTGRGSPSEAMHLGHVLPFTTKLQDVFNVL	195			
Qy	122	VIQMTDEKYLKMDLLDQAYGDAVENAKDIIACGFDINTKFIISDLDYMGSSGFYKVV	181			
Db	196	VIQMSDEKYLKMDLLQAYSYVENAKDIIACGFDINTKFIISDLEWYGQSPGFYKVV	255			
Qy	182	VKIQKATVFNQVKIFSGFTSDCGIKISFPAIQAAFPSSNSFPQIFRPTDIOCLIPCAI	241			
Db	256	VKIQKATVFNQVKIFGFTSDCGIKISFPAVQAAFPSSNSFPKIFRPTDIOCLIPCAI	315			
Qy	242	DQDEYFMTBDAVARIGYKPKPALHSTFPFPAQQAOTMSASDPNSSIFLDTAKQIKTK	301			
Db	316	DQDEYFMTBDAVARIGHKPKPALHSTFPFPAQQAOTMSASDPNSSIFLDTAKQIKTK	375			
Qy	302	VNKIAFGSGRDTIEHRQFGNCDCVDVSPMYLTFPLEDDDKLEJOIRKDYTGAMLTGELK	361			
Db	376	VNKIAFGSGRDTIEHRQFGNCCEVDVSPMYLTFPLEDDDKLEJOIRKDYTGAMLTGELK	435			
Qy	362	KALIEVLQPLIAEHQARKKETEIDIEVEFMPRLSTDFQ	401			
Db	436	KTLIDIVQPLIAEHQARKKATEIDIEVEFMPRLQSTDFQ	475			

RESULT 3	Q9D8R9	PRELIMINARY;	PRT;	329 AA.
ID	Q9D8R9			
AC	Q9D8R9;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Trypophanyl-tRNA synthetase.			
GN	WARS.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/60; TISSUE=PANCREAS;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Aachui J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,			
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kaenkawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischnann W., Gaasterland T., Gissi C., King B., Kochiya H.,			
RA	Knehl P., Lewis S., Matsuo Y., Nikaishi I., Pesole G., Quackenbush J.,			
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bersh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bernaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita A., Gariboldi M.,			
RA	Guistincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Maehama J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald L., Maehama J., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontuski S.,			

RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK007754; BAB25235.1; -.  
 DR MGD; MGI:104630; Wats.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; Ttp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244C4A2 CRC64;

Query Match 77.0%; Score 1629; DB 11; Length 329;  
 Best Local Similarity 92.1%; Pred. No. 2.7e-131;  
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 73 MNQVLDAYENKKPFLYLTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIOQTDEKYL 132  
 DB 1 MNQVLDAYENKKPFLYLTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIOQTDEKYL 60  
 QY 133 WKDLTLDQAYGDAVENAKIILACGPDINTFTFSDLDYMGSSGFYKVVKI QKHTVNO 192  
 DB 61 WKDLTLEQAYSYTVENAKIILACGPDINTFTFSDLDYMGSSGFYKVVKI QKHTVNO 120  
 QY 193 VKGIFGTFSDICIGKISFPALQAPSFNSPQIFPRDRDIOCLIPCAIDODPYFPMTRD 252  
 DB 121 VKGIFGTFSDICIGKISFPALQAPSFNSPQIFPRDRDIOCLIPCAIDODPYFPMTRD 180  
 QY 253 VAPRIGYPRPALHSTFFPALQAGQTKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGSD 312  
 DB 181 VAPRIGYPRPALHSTFFPALQAGQTKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGSD 240  
 QY 313 TIEEHROFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGELKKALILEVLOPLI 372  
 DB 241 TIEEHROFGNCDDVDSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGELKKALILEVLOPLI 300  
 QY 373 AEHQARRKEVTDEIYKEFTMPKRLSFDQ 401  
 DB 301 AEHQARRKEVTDEIYKEFTMPKRLSFDQ 329

## RESULT 4

QY 070184 PRELIMINARY; PRT; 305 AA.  
 AC 070184;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Trypophanyl-tRNA synthetase (Fragment).  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HARTLEY; TISSUE=SPLEEN;  
 RA Yang D., Goto K., Watanabe N., Kobayashi Y.;  
 RT "Identification and Cloning of Genes Whose Expressions are Elevated  
 RT during DNCB-induced Guinea Pig Skin Delayed-type Hypersensitivity  
 RT Reaction."  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB012222; BAA25288.1; -.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR002306; Ttp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR Aminoacyl-tRNA synthetase.  
 KW NON TER 1  
 FT SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 72.6%; Score 1537; DB 11; Length 305;  
 Best Local Similarity 94.4%; Pred. No. 1.8e-123;

Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 97 EAMHVGHLIPFIETKMLQDVENVPLVIOQTDEKYLWKDLTLDQAYGDAVENAKIILACG 156  
 DB 1 EAMHVGHLIPFIETKMLQDVENVPLVIOQTDEKYLWKDLTLDQAYGDAVENAKIILACG 60  
 QY 157 PINTNFIETSDLDYMGSSGFYKVVKI QKHTVNOVGIKFTSDICIGKISFPALQAA 216  
 DB 61 PINTNFIETSDLDYMGSSGFYKVVKI QKHTVNOVGIKFTSDICIGKISFPALQAA 120  
 QY 217 PSFNSPQIFPRDRDIOCLIPCAIDODPYFPMTRDVAIPRIGYPRPALHSTFFPALQGA 276  
 DB 121 PSFNSPQIFPRDRDIOCLIPCAIDODPYFPMTRDVAIPRIGYPRPALHSTFFPALQGA 180  
 QY 277 QTKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGSDTIEEHROFGNCDDVDSFMYLTFF 336  
 DB 181 QTKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGSDTIEEHROFGNCDDVDSFMYLTFF 240  
 QY 337 LEDDDKLEQIRKDYTSGAMLTGELKKALILEVLOPLIAEHQARRKEVTDEIYKEFTMPKRL 396  
 DB 241 LEDDDKLEQIRKDYTSGAMLTGELKKALILEVLOPLIAEHQARRKEVTDEIYKEFTMPKRL 300  
 QY 397 SFDQ 401  
 DB 301 SFDQ 305

## RESULT 5

QY 09U4Y0 PRELIMINARY; PRT; 420 AA.  
 AC 09U4Y0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Trypophanyl-tRNA synthetase (Fragment).  
 GN AATS-TRP OR CG9735.  
 GN Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99250164; PubMed=10233165;  
 RA Sennatiah P., Andrew D.O.;  
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in  
 RT the developing Drosophila salivary gland."  
 RL Mol. Biol. Cell 10:1595-1608(1999).  
 DR EMBL; AF125157; AAF20167.1; -.  
 DR FlyBase; FBgn010803; Aats-trp.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; Ttp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.  
 FT NON TER 1  
 SQ SEQUENCE 420 AA; 46975 MW; 4AF7088AF426AEB6 CRC64;

Query Match 65.0%; Score 1374.5; DB 5; Length 420;  
 Best Local Similarity 64.8%; Pred. No. 2.4e-109;  
 Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;

QY 7 ATEAEEDVDDPMYVOTSSAKGIDYDKLIVRFSSKIDDELINRIERATGQRPHHFLRRGI 66  
 DB 28 ATEAEEDVDDPMYVOTSSAKGIDYDKLIVRFSSKIDDELINRIERATGQRPHHFLRRGI 87  
 QY 67 FFSSHDMMQVLDAYENKKPFLYLTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIOQT 126  
 DB 88 FFSSHDMLTILTLBQKGFYLTGRGSSSEAMHVGHLIPFIETKMLQDVENVPLVIOQT 147

```

QY 127 DDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKVVYIOK 186
DB 148 DDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKVVYIOK 207
QY 187 HTFENOVKGIFFGTDSDICGISFPALQOAPSFNSFPQIFRDRDIOCLIPALIDDPY 246
DB 208 CVTFENOVKGIFFGTDSDICGISFPALQOAPSFNSFPQIFRDRDIOCLIPALIDDPY 266
QY 247 FRMTDVAPRIGYIPKALHSTFFPALQOAGTKMSADPNSSIFLDTAOKITKYNKA 306
DB 267 FRMTDVAPRIGYIPKALHSTFFPALQOAGTKMSADPNSSIFLDTAOKITKYNKA 326
QY 307 FSGGRDTEHROFGNCVDVSMYLTFFLEDDDKLEOIRKDYTGAMLTGELKALIE 366
DB 327 FSGGRVTVHEHRKLGCVPEVDVSYOLKFFLEDDDKLEVRVAVSKGEMLTGELKALIE 386
QY 367 VLQPLAEHQARKEVTEIVKSFMPRKLSP 398
DB 387 TLTPIVEHQARKEVTEIVKSFMPRKLSP 418

```

## RESULT 6

Q904Y1 PRELIMINARY; PRT; 430 AA.

```

ID 0904Y1
AC 0904Y1;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Trypophanyl-tRNA synthetase (At21437p).
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.;
RX MEDLINE=99250164; Pubmed=10233165;
RA Seshiah P., Andrew D.J.;
RT "WRS-85b: A tryptophanyl-tRNA synthetase expressed to high levels in
  the developing Drosophila salivary gland."
RL Mol. Biol. Cell 10:1595-1608(1999).
RN 12
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez C., Guarini H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Munoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF125156; AA020166.1; -
DR EMBL; AY075249; AL68116.1; -
DR FlyBase; FBgn0010803; Aats-trp.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASINTTRP.
DR TIGR; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA LIGASE_1; 1.
KW Aminoacyl-tRNA synthetase.
SQ SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;

```

Query Match 65.0%; Score 1374.5; DB 5; Length 430;  
 Best Local Similarity 64.8%; Pred. No. 2.5e-109;  
 Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;

```

QY 7 ATAAEDPDVPMVYVSSAKGIDYDKIYFSGSSKIDKELINIRATGQRPHHPLRRGI 66
DB 38 AATPTEVDVPMVYVSSAKGIDYDKIYFSGSSKIDKELINIRATGQRPHHPLRRGI 97
QY 67 FFSHRDMQVLDAYENKPPYLYTGRGSSSEAMVGHLLPIFTKMLQDVFNPLVYIOMT 126

```

```

DB 98 FFSHRDLHLITLRECKGFYLYTGRGSSSGSLHGHVLPFIMTKLOETFDVPLVYIOLT 157
QY 127 DDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKVVYIOK 186
DB 158 DDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKFIFSDLDYMGMSGGFYKVVYIOK 217
QY 187 HTFENOVKGIFFGTDSDICGISFPALQOAPSFNSFPQIFRDRDIOCLIPALIDDPY 246
DB 218 CVTFENOVKGIFFGTDSDICGISFPALQOAPSFNSFPQIFRDRDIOCLIPALIDDPY 276
QY 247 FRMTDVAPRIGYIPKALHSTFFPALQOAGTKMSADPNSSIFLDTAOKITKYNKA 306
DB 277 FRMTDVAPRIGYIPKALHSTFFPALQOAGTKMSADPNSSIFLDTAOKITKYNKA 336
QY 307 FSGGRDTEHROFGNCVDVSMYLTFFLEDDDKLEOIRKDYTGAMLTGELKALIE 366
DB 337 FSGGRVTVHEHRKLGCVPEVDVSYOLKFFLEDDDKLEVRVAVSKGEMLTGELKALIE 396
QY 367 VLQPLAEHQARKEVTEIVKSFMPRKLSP 398
DB 397 TLTPIVEHQARKEVTEIVKSFMPRKLSP 428

```

## RESULT 7

Q9VHG2 PRELIMINARY; PRT; 430 AA.

```

ID 09VHG2
AC 09VHG2;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Aats-trp protein.
GN AATS-TRP OR CG9735.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RC MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandlis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miles G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benise P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M.C., Casale S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Galbati W.M., Glasser K.,
RA Gosler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulj D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshireli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```





```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99065613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL132853; CAB60439.1; -.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR00579; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
SQ SEQUENCE 324 AA; 36289 MW; 6E687E6E420ECC12 CRC64;

Query Match 48.5%; Score 1027; DB 5; Length 324;
Best Local Similarity 64.9%; Pred. No. 9.6e-80;
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

QY 4 GPDATAEEDPVPWTVQTSASAKIDYDKLIVFGSSKIDKELINIRERATGQRPHHFLR 63
DB 23 GGGVQDEDEBRVTPWEVTTTKATGIDYDKLIVFGCKLDEILARERTGCHKASPMR 82
QY 64 RGIFFSHRDMNOVLDAVENKKPFLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVI 123
DB 83 RGMFFAHRDLTALIDRKEQGRPFYLYTGRGASSGSLHGLVFPFIETKMLQEVFDPVLVI 142
QY 124 QMTDDEKYLKMDLTDOAVGDVAENAKDIIACGPDINKETIFSDLDVMGSSGFYKNVVK 183
DB 143 QMTDDEKYLKMDLVDAENAKDIIACGPDINKETIFSDLDVMGSSGFYKNVVK 200
QY 184 IQHVTENQYKGIFFGTDSDICIGISFPALQAAAPSFNSPQIFRDRDIOCLIPCAIDQ 243
DB 201 IWKVNTNQARALFGFTPEDCGKAAPVAAAPCFASSPQIFGKNDIPCLIPCAIDQ 260
QY 244 DPFYFNRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKV 302
DB 261 DPFYFNRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKV 319

RESULT 10
Q8SOY5 PRELIMINARY; PRT; 385 AA.
AC Q8SOY5;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Trypophanyl tRNA synthetase.
GN EC011 0530.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
OX NCBI_TaxID=6035;
QY RN [1] SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2] SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prenier G., Barbe V., Peyretailade E., Brotier P., Wincker P.,
RA Delbec F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weisenbach J., Vives C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi."
RL Nature 414:450-453(2001).
DR EMBL: AL590450; CAD25963.1; -.
SQ SEQUENCE 385 AA; 44188 MW; 2865C3AC95FCC859 CRC64;

Query Match 44.5%; Score 942; DB 5; Length 385;
Best Local Similarity 46.1%; Pred. No. 2.3e-72;
Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

```

```

QY 12 EDFVDPMTVQTSASAK-----GIDYDKLIVFGSSKIDKELINIRERATGQRPHHFLR 67
DB 3 EQRTIPMDVEAVYSTDEVPALIDYDKLINFGCKEPQALADLEKLSGPAHYFFRRGIV 62
QY 68 FSHRDMNOVLDAVENKKPFLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVIQMTD 127
DB 63 FAHRDNLFLDLEIANRPFYLYTGRGSSKTHIGHITIFPLCKWQDAFKRLVLIQTD 122
QY 128 DEKYLKMDLTDOAVGDVAENAKDIIACGPDINKETIFSDLDVMGSSGFYKNVVK 187
DB 123 DEKYLKMDLTDOAVGDVAENAKDIIACGPDINKETIFSDLDVMGSSGFYKNVVK 179
QY 188 VTFNQYKGIFFGTDSDICIGISFPALQAAAPSFNSPQIFRDRDIOCLIPCAIDDPYF 247
DB 180 ININEAIKFGFDMSNIGQVGPFAKEIAPCCSSSRFIFGK---GAMCLVPAVADDPFF 236
QY 248 RMTRDVAPRIGYKPKPALHSTFPALQAGQTKMSADPNSSIFLDTAQIKTKVKNKAF 307
DB 237 RLARDKAKALGKKKPSIYVSLPDLKGVNRKMSADPNSSIFLDTAQIKTKVKNKAF 296
QY 308 SGGRTIEHRQPGNCVDVSGFMVITFLEDDKLEQIRKYTSGAMLTGELKALIV 367
DB 297 SGGRTIEHRQPGNCVDVSGFMVITFLEDDKLEQIRKYTSGAMLTGELKALIV 356
QY 368 LOPLIAHOARREVTDEIVKEFMTPRK 395
DB 357 IQEFVSRVQESKRKRTDLDLRAFDINK 384

RESULT 11
Q8U453 PRELIMINARY; PRT; 385 AA.
AC Q8U453;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Trypophanyl tRNA synthetase.
GN PF0241.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
QY RN [1] SEQUENCE FROM N.A.
RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.W., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE010149; AL80365.1; -.
KW Aminocyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 385 AA; 45178 MW; 3A7A628958200CCC CRC64;

Query Match 39.1%; Score 826.5; DB 17; Length 385;
Best Local Similarity 45.9%; Pred. No. 1.8e-62;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

QY 11 EEDF-VDPMTVQTSASAKIDYDKLIVFGSSKIDKELINIRERATGQRPHHFLR 69
DB 2 EEFKVTPEVEGII---IDYKLLIEQFTSPPLTDLLEKRLAKLISELPFRRKFFFS 57
QY 70 HRDMNOVLDAVENKKPFLYTGRCPSSEAMVGHLPFTKMLQDVFNPLVIQMTD 129
DB 58 HRDVKVLDVDEQKGFYLYTGRGSG-PMHGHIIIPFATKMLDEKFPVNLVIQITD 116
QY 130 KYLWMD-LTLDQAVGDVAENAKDIIACGPDINKETIFSDLDVMGSSGFYKNVVK 188
DB 117 KFLFKNELTFEEDTKYWAYQNIIDIIAVGDPDKTFIFQSSF---TKIYENAIPIAKKI 172
QY 189 TFNQYKGIFFGTDSDICIGISFPALQAAAPSFNSPQIFRDRDIOCLIPCAIDDPYR 248
DB 173 NFSMAVAVGFTQESKIGMIFFPALQAAAPF-----FEKK---RCLIPAAIDDPYR 222

```

QY 249 MTRDVAPRIGYKPKPALHSTFPALOGAOTKMSASDPNSSIFLPTAKOIKTKVKNKHAAS 308  
 DB 223 LORDABESIGYKTKALHSTFPALLEGKMSASDPNFAITLTPNPEAGKIKKFAIT 282  
 QY 309 GGRDITIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDY---TSGMLTGEELKALI 365  
 DB 283 GGQPLTEKREKGNPEKCVFPMLEIFEPDDK--KLMEYVACWNGELGCECKRYLI 340  
 QY 366 EYLOPLIAEHQARKEVDEIVKEFMTPKLS 397  
 DB 341 QVQEFLEKHEORRKK-AEKLVEKFKYTKLA 371

## RESULT 12

Q9UIF5 PRELIMINARY; PRT; 490 AA.  
 ID Q9UIF5  
 AC Q9UIF5;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase.  
 GN L1063.04.  
 OS Leishmania major.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRIDELIN;  
 RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,  
 RA Barrett B.G.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRIDELIN;  
 RA MEDLINE=98146435; PubMed=9477341;  
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 RA Smith D.F.;  
 RL "A physical map of the Leishmania major Friedlin genome."  
 RL Genome Res. 8:135-145(1998).  
 DR EMBL; AL121862; CAB58393.1;  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; Ttp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRANSSYNTHTRP.  
 DR PROSITE; PS00178; AA\_tRNA\_LIGASE\_I; 1.  
 DR Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 490 AA; 54904 MW; 172C68622C511D3D CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;  
 Best Local Similarity 38.1%; Pred. No. 1.2e-59;  
 Matches 173; Conservative 87; Mismatches 129; Indels 65; Gaps 11;

QY 3 HGPDATEAEEDFVDPWTQTSAGKIDYDKLIVFGSSKID-----KELINRIERAT- 54  
 DB 40 HPRDGDAGAEDEVITPVVAAKGGQGINVRLTIFAEKMDGARGHMDGVAKCKKRM 99  
 QY 55 -----GQR-----HFLRGIFFSHDM 73  
 DB 100 TTTPSANAEGVATPVLDAAVAGDALQAASAPMOCROOPPALALHFFHRDIASFHRDL 159  
 QY 74 NGUVAAYE-----NKKPFYLYTGRGSSSEAMVGHILPIFTFKMLQDVNVPLVIOMTD 128  
 DB 160 HRAALVDIESVKTGESVFLYTGKSGAGTMHGHVLPMLTKYLDVFSLPVLIQITD 219  
 QY 129 ERYLWKDLTLDOAYGDAV--ENAKDIIACGFDINKTFISDLDYMGSSGFYKXNVKIQK 186  
 DB 220 EKFLPRDVPFEAGKADILRSNIKOILAFNPNRHFIFIRNTHYMG---DMVPTVRLQR 276  
 QY 187 HTYENOVKGIFGTSDDCIGKISFPALQAAFSFNSFPQIFR--DRTDIQCLIPCAIDDD 244  
 DB 277 SMTGNANVKTGLITGTDSDNVKGLAFPATOAFCPSTAFRVLNRGDR-PKRCILPCAIDDD 335

QY 245 PYFRMTRDVAPRIGYKPKPALHSTFPALOGAOTKMSAS-DPNSSIFLPTAKOIKTKVKN 303  
 DB 336 PPFUTLRAALRLKLPALHSTFKLPALKGLHHMSSSABEKVITLHDDKQVRKL- 394  
 QY 304 KHAESGGRDITIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKA 363  
 DB 395 BRAFGGCGATLEQOMQETGANIELDVAOYLRFPCDDTLFADVTQYRSGLTNSGEVDL 454  
 QY 364 LIE-VLOPLIAEHQARKEVDEIVKEFMTPKL 396  
 DB 455 AADCIIRVLIHWRERRATVTDVVEECRIADI 488

## RESULT 13

Q95295 PRELIMINARY; PRT; 136 AA.  
 ID Q95295  
 AC Q95295;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Tryptophanyl-tRNA synthase (Fragment).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SMALL INTESTINE;  
 RA Winzore A.K., Fredholm M., Davies W.;  
 RL "Evaluation and characterization of a porcine small intestine cDNA library."  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z81267; CAB03585.1;  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR002306; Ttp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRANSSYNTHTRP.  
 FT NON\_TER 1 1  
 FT NON\_TER 136 136  
 SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

Query Match 32.2%; Score 682; DB 6; Length 136;  
 Best Local Similarity 94.1%; Pred. No. 1e-50;  
 Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 211 PAIQAAFSFNSFPQIFDRDIDQCLIPCAIDDDPYFRMTRDVAPRIGYKPKALHSTFF 270  
 DB 1 PAIQAAFSFNSFPQIFDRDIDQCLIPCAIDDDPYFRMTRDVAPRIGYKPKALHSTFF 60  
 QY 271 PALOGAOTKMSASDPNSSIFLPTAKOIKTKVKNKHAESGGRDITIEHROFGNCDVVSF 330  
 DB 61 PALOGAOTKMSKXANSSIFLPTAKOIKTKVKNKHAESGGRDITIEHROFGNCDVVSF 120  
 QY 331 MYLTFLEDDDKLEQI 346  
 DB 121 MYLTFLEDDDKXEQI 136

## RESULT 14

O8TYF7 PRELIMINARY; PRT; 374 AA.  
 ID O8TYF7  
 AC O8TYF7;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase.  
 GN TRPS OR MK0343.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE=21927647; PubMed=11930014;  
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polunin N.N.,  
RA Shcherbina O.V., Shakova V.V., Belova G.I., Artynd L.,  
RA Natala D.A., Rogozin I.B., Ratsov R.L., Wolf Y.I., Seltzer K.O.,  
RA Mal'kh A.G., Koonin E.V., Koz'yavkin S.A.,  
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
RT and monophyly of archaeal methanogens.";   
Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
DR EMBL: AE010331; AAM01558.1; -;  
KM Aminocacyl-tRNA synthetase, Complete Proteome  
SQ SEQUENCE 374 AA; 4273 MW; 0BD4EB579321E828 CRC64;

[illegible]

RESULT 15	ID	QDTUT1	PRELIMINARY;	PRT;	437 AA.
AC	QDTUT1;				
DT	01-JUN-2002 (TREMBlrel. 21, Created)				
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)				
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)				
DE	TrypTopophanyl-tRNA synthetase.				
CS	TRPS OR MA0172.				
OC	Methanosarcina acetivorans.				
OC	Archaea; Euryarchaeota; Methanococci; Methanosarcinales;				
OC	Methanosarcinaceae; Methanosarcina.				
ON	NCBI_TaxID=2214;				
XX	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C2A / ATCC 35395 / DSM 2834;				
RX	MEDLINE=21929760; PubMed=11932228;				
RA	Galagan J.E., Nussbaum C., Roy A., Ferriszi M.G., Macdonald P.,				
RA	FitzHugh W., Calvo S., Engels R., Smirnov S., Anrood D., Brown A.,				
RA	Allen N., Naylor J., Stange-Thomann N., DeRellano K., Johnson R.,				
RA	Linton L., McEwen P., McKernan K., Talamas J., Tittrell A., Ye W.,				
RA	Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Gus A.M.,				
RA	Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,				
RA	Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,				
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,				
RA	Perry J.G., Jarrell K.F., Uing H., Macario A.J.L., Paulsen I.,				
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,				

RA Metcalf W.W., Birren B.  
RT "The genome of *Methanosaeta acetivorans* reveals extensive metabolic  
RT and physiological diversity.",  
RL Genome Res. 12:532-542 (2002).  
DR EMBL: AEO10675; AAM03625.1; -.  
KY Antinucy -tRNA synthetase; Complete proteome.  
SQ SEQUENCE 437 AA; 45067 MW; 10DF8FBBBF7A65 CRC64;

Query Match	15.6%	Score 329.5	DB 17	Length 437
Best Local Similarity	25.0%	Pred. No. 8.1e-20		
Matches 113	Conservative 67	Mismatches 171	Indels 101	Gaps 14

  

QY	15	VDPMTVYSSANGIDYDKLIVRFGSSKIDKELINRIRATGQRPHNLRGCFIFSRDPA	74
Db	5	LDPW-----SSSDITYSKLFEEFGISPEENVL-----PEIPSHMYMRKVLFGHRDY	54
QY	75	QVLDVAENKKPYLTGRGSPSEAMHWCHLIPFIPTKLDPDVFVPLVIMQTDKELMK	134
Db	55	QIAEMRTGAPSPVNDGFMPSGK-VHLCHKVMQIYHDEM-GASAFVGIARREAFSVR	112
QY	135	DLTLDAVDAVEN-AKDI IACGPDINKTIFSLDLVMGSSGFYKNV-KIQCHVTENO	192
Db	113	GFSGWQCREIGVEEYILSLIALGFEPD-----GLIYFSGGCSVKKLAFELCAKNFSE	166
QY	193	VKGIPEFTDSDCIKISFPALQAPSPFSNPFOIFRDRTDIQCLIPCAIMQDYPFMTD	252
Db	167	LSAIIYFSGSETSLSHMLSVATQADILQ--PQLVEEFGKFPVVVPGPDPHLALTGG	223
QY	253	VAPRIQ-----YP- 260	
Db	224	LAKGNMFPVEEREDVTKGRKYLSVRGKTAQKEALQELKRIIPGKVLVEEHIIDVLEYP	283
QY	261	-----KPAIIHSTFPALQAGQTKMSASDPNPSIFLDTPAKQIK 299	
Db	284	LAGELETVRETVVEEGCAFIIPPASTYHRFMSGIQQG--KMSSSIIPSQIALDLSRKEGA	341
QY	300	TKVNKHAIFSGBDTIEEHROFGNCDVDSFMVLTFFL-BDDDKLEQIRKDYTSGAMLTG	358
Db	342	KKV-KRAKGGCVTLEBQKLLGKRECEVSVELMALPHILASDELEIKQECISGTRMGQ	400
QY	359	ELKKALIEVLOPLIAEHQARKEVTDVKEF 390	
Db	401	SCKQIAAEHQEFLKDHOEKR-ELAREHLDIEY 431	

Search completed: July 10, 2003, 12:31:41  
Job time : 28.2194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:26:29 ; Search time 8.74909 Seconds

(without alignments)  
1900.998 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SMHGDATEAEEDFVDPWTV.....VNDIEVKEFMTPRKLSFDDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	99.3	471	1	SYN_HUMAN
2	2028.5	95.9	475	1	SYN_BOVIN
3	1961.5	92.7	475	1	SYN_RABIT
4	1938	91.6	481	1	SYN_MOUSE
5	1210	57.2	395	1	SYN_SCHPO
6	1163	55.0	432	1	SYN_YEAST
7	910.5	43.0	381	1	SYN_SULTO
8	905	42.8	380	1	SYN_SULSO
9	803	37.9	385	1	SYN_PYRAB
10	796	37.6	386	1	SYN_PYRHO
11	735	34.7	375	1	SYN_PYRAE
12	534.5	25.3	380	1	SYN_HALNI
13	409.5	19.4	370	1	SYN_METJA
14	397.5	18.8	364	1	SYN_METHA
15	386	18.2	134	1	SYN_ENCCU
16	370.5	17.5	420	1	SYN_ARCFU
17	353	16.7	374	1	SYN_AERPE
18	299.5	14.2	426	1	SYN_THEVO
19	268.5	12.7	426	1	SYN_TREBA
20	192	9.1	323	1	SYN_ARCFU
21	185.5	8.8	341	1	SYN_STRPN
22	182.5	8.6	340	1	SYN_STRPY
23	175.5	8.3	341	1	SYN_CLODO
24	174.5	8.2	341	1	SYN_LACIA
25	172	8.1	394	1	SYN_YEAST
26	170.5	8.1	366	1	SYN_SUSO
27	169	8.0	395	1	SYN_AQUIA
28	163.5	7.9	346	1	SYN_AQLTR
29	163.5	7.7	353	1	SYN_BORBU
30	159.5	7.5	337	1	SYN_TREPA
31	158	7.2	346	1	SYN_CHLPU
32	152	7.2	344	1	SYN_CHLPPN
33	146	6.9	326	1	SYN_HELPY

34	138	6.5	319	1	SYN_METHH	027795 methanobact
35	136.5	6.5	528	1	SYN_BOVIN	029465 bos taurus
36	135.5	6.4	328	1	SYN_THEMA	094972 thermotoga
37	133	6.3	328	1	SYN_BACST	P00953 bacillus st
38	132.5	6.3	343	1	SYN_MYCLE	Q49901 mycobacteri
39	132.5	6.3	528	1	SYN_HUMAN	P54577 homo sapien
40	130	6.1	326	1	SYN_HELPJ	0921x4 hellicobact
41	129	6.1	401	1	SYN_SCHPO	014055 schizosacch
42	128.5	6.1	379	1	SYN_YEAST	P04803 saccharomyc
43	128	6.0	334	1	SYN_HAEIN	P43835 haemophilus
44	124.5	5.9	347	1	SYN_MYCGE	P47372 mycoplasma
45	122	5.8	350	1	SYN_CAEBL	P46579 caenorhabdi

## ALIGNMENTS

RESULT 1	SYN_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381; P78535; Q9UDJ3;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)			
DE	(TrpRS) (TrpP53) (hMRS).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92105071; PubMed=1761529;			
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92107982; PubMed=1763065;			
RA	Flechner J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92112058; PubMed=1765274;			
RA	Frolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovieva O.L.,			
RA	Kiselev L.V.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92164636; PubMed=1537332;			
RA	Buwallt U., Flohr T., Boettger B.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RN	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RX	MEDLINE=96319994; PubMed=8724762;			
RA	Sokolova I.V., Narovianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-			
RT	tRNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RN	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE=Spem;			
RX	MEDLINE=93292992; PubMed=7685728;			

RA Prolova L.Y., Grigorieva A.Y., Sudomoina M.A., Kisselev L.L.;  
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-  
 response elements and exon-intron organization.";  
 RL Gene 128:237-245(1993).  
 RN [7]  
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.  
 RC TISSUE=Keratinocytes; PubMed=1286667;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Collis J.E.,  
 Vandekeckevoete J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 RN [8]  
 RP FUNCTION  
 RX MEDLINE=9225128; PubMed=1373391;  
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;  
 RT "An interferon-induced protein with release factor activity is a  
 tryptophanyl-tRNA synthetase.";  
 RL FEBS Lett. 300:162-166(1992).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- INDUCTION: BY INTERFERON GAMMA.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC EMBL; M77804; AAA67324.1; -  
 DR EMBL; X59892; CAA42545.1; -  
 DR EMBL; M61715; AAA61298.1; -  
 DR EMBL; X62570; CAA44450.1; -  
 DR EMBL; S82905; AAB39381.1; -  
 DR EMBL; X67920; CAB94198.1; -  
 DR EMBL; X67921; CAB94198.1; JOINED.  
 DR EMBL; X67922; CAB94198.1; JOINED.  
 DR EMBL; X67923; CAB94199.1; -  
 DR EMBL; X67924; CAB94199.1; JOINED.  
 DR EMBL; X67925; CAB94199.1; JOINED.  
 DR EMBL; X67926; CAB94199.1; JOINED.  
 DR EMBL; X67927; CAB94199.1; JOINED.  
 DR EMBL; X67928; CAB94199.1; JOINED.  
 DR PIR; A41706; A41706.  
 DR PIR; A41633; A41633.  
 DR PIR; JH0533; JH0533.  
 DR PIR; S19246; S19246.  
 DR Aarhus/Ghent-2DPAGE; 3524; IEF.  
 DR PHCI-2DPAGE; P23381; -  
 DR Genew; HGNC:12729; WARS.  
 DR MIM; 191050; -  
 DR InterPro; IPR002306; Trp-tRNA-synt\_1b.  
 DR InterPro; IPR000738; WHEP-TRS.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA tRNA LIGASE\_I; 1.  
 DR PROSITE; PS00762; WHEP-TRS; 1.  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT DOMAIN 19 64 WHEP-TRS  
 FT SITE 164 173 "HIGH" REGION.  
 FT SITE 349 353 "KMSKS" REGION.  
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
 FT CONFLICT 424 424 A -> R (IN REF. 4).

SQ SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;  
 Query Match 99.3%; Score 2101; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 2,8e-165;  
 Matches 399; Conservative 0; Mismatch 2; Indels 0; Gaps 0;  
 QY 1 SNHGPGATEAEEDFVDPMTVQTSSAKGIDYDKLIYVFGSSKIDKELINERATGQRPNN 60  
 DB 71 SNHGPGATEAEEDFVDPMTVQTSSAKGIDYDKLIYVFGSSKIDKELINERATGQRPNN 130  
 QY 61 FLRGGIFFSHRDNNQVLDAVENKPPYLYTGRGSSSEAMVGHILPFTKMLQDVFNVP 120  
 DB 131 FLRGGIFFSHRDNNQVLDAVENKPPYLYTGRGSSSEAMVGHILPFTKMLQDVFNVP 190  
 QY 121 LVIQMTDDERKYLKMDLTLLQAVGDAVENAKDIIACGFDINKTTFISDDLVMGSSGFYRN 180  
 DB 191 LVIQMTDDERKYLKMDLTLLQAVGDAVENAKDIIACGFDINKTTFISDDLVMGSSGFYRN 250  
 QY 181 VKKIQKHTFNVQKGIFFGFTSDSDICKISFPALQAPSPNSPFOIFRRTDIQCLIPCA 240  
 DB 251 VKKIQKHTFNVQKGIFFGFTSDSDICKISFPALQAPSPNSPFOIFRRTDIQCLIPCA 310  
 QY 241 IDDPYFRMTDVAAPRIGYKPKALHSTFPFALQAGTQMSASDPNSSIFLDTAKQIKT 300  
 DB 311 IDDPYFRMTDVAAPRIGYKPKALHSTFPFALQAGTQMSASDPNSSIFLDTAKQIKT 370  
 QY 301 KVNKAHFGSGRDTIEHROFGNCDVDVSPMLTFLEDDDKLEQRKYTGAMLTGEL 360  
 DB 371 KVNKAHFGSGRDTIEHROFGNCDVDVSPMLTFLEDDDKLEQRKYTGAMLTGEL 430  
 QY 361 KKALIEVLOPLAEHQARKEVTDIVKEPMTPRKLSFPFO 401  
 DB 431 KKALIEVLOPLAEHQARKEVTDIVKEPMTPRKLSFPFO 471  
 RESULT 2  
 SYM\_BOVIN  
 ID SYM\_BOVIN STANDARD; PRT; 475 AA.  
 AC P17248;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (Trps).  
 GN WARS.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina;  
 RX MEDLINE=91329348; PubMed=1907847;  
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,  
 Gandar J.-C., Benedetto J.-P., Sallafraque M.-L., Alterio J.,  
 Gueguen M., Sargier C., Labouesse B., Bonnet J.;  
 RA "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 prokaryotic synthetases but near identity with mammalian peptide  
 chain release factor.";  
 RT Biochemistry 30:7809-7817(1991).  
 RL [2]  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,  
 Gueguen M., Benedetto J.-P., Sargier C., Alterio J., La Bouesse B.,  
 Labouesse J., Bonnet J.;  
 RA Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 RL -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: X53918; CAA37872.1; -
DR EMBL: X52113; CAA36356.1; -
DR PIR: A40279; YWBO.
DR PIR: S14540; S14540.
DR InterPro: IPR002306; TTP-trna-synt_1b.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; trna-synt_1b.
DR InterPro: IPR001412; trna-synt_1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR Pfam: PF00579; trna-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR TIGRPFAM: TIGR00233; trps; 1.
DR PROSITE: PS00178; AA-trna-ligase_1; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
DR KMW: Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 24 69 WHEP-TRS.
FT SITE 169 178 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT CONFLICT 17 17 L -> M (in REF. 2).
SQ SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 95.9%; Score 2028.5; DB 1; Length 475;
Best Local Similarity 95.8%; Pred. No. 2.6e-159;
Matches 364; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SHNGPDTEAEEDFDVPMVTQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPH 60
DB 76 SGEGLDATEADEDFVDPMTVTQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPH 135
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
DB 136 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 195
QY 121 LVIOQMTDEKYLKQDLTDOAYGDAVENKDIAGSPDINKTFISDLDYMGSSGFYKN 180
DB 196 LVIOQMTDEKYLKQDLTDOAYGDAVENKDIAGSPDINKTFISDLDYMGSSGFYKN 254
QY 181 VKIQRHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 240
DB 255 VKIQRHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 314
QY 241 IQDDYFRMTRVAPRIGYPRKPALHSTFPPALQAGQTMASADPNSSIFLDTAKQIKT 300
DB 315 IQDDYFRMTRVAPRIGYPRKPALHSTFPPALQAGQTMASADPNSSIFLDTAKQIKT 374
QY 301 KKNKAFFSGRPTIEHROFGNCVDVSPMYLTFLEDDDKLEQRKRYTGAMLTGEL 360
DB 375 KKNKAFFSGRPTIEHROFGNCVDVSPMYLTFLEDDDKLEQRKRYTGAMLTGEL 434
QY 361 KKALEVLQPLIAEHQARRKEVTEIVKSEFMTPRKLSFDFQ 401
DB 435 KKALEVLQPLIAEHQARRKEVTEIVKSEFMTPRKLSFDFQ 475

RESULT 3
SYN_RABIT
ID SYN_RABIT STANDARD; PRT; 475 AA.
AC P23612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-NOV-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).

```

```

GN WARS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigen W.J., Murny D.M., Harlow E., Caeky C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases."
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RX MEDLINE=9409008; PubMed=8404687.
RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,
RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins."
RL EMBO J. 12:4013-4019(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(TTP) = AMP +
CC dihydrophate + L-tryptophanyl-tRNA(TTP).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTRAINS 1 WHEP-TRS DOMAIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: M33460; AAA31246.1; ALT_SEQ.
DR EMBL: U02595; AAB60257.1; -.
DR PIR: A35904; YMRBPR.
DR InterPro: IPR002306; TTP-trna-synt_1b.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; trna-synt_1b.
DR InterPro: IPR001412; trna-synt_1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR Pfam: PF00579; trna-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR TIGRPFAM: TIGR00233; trps; 1.
DR PROSITE: PS00178; AA-trna-ligase_1; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
DR KMW: Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
SQ SEQUENCE 475 AA; 53729 MW; 33BC9E718FFA5DC4 CRC64;

Query Match 92.7%; Score 1961.5; DB 1; Length 475;
Best Local Similarity 91.5%; Pred. No. 8.5e-154;
Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 2 NHG-PDTEAEEDFDVPMVTQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPH 60
DB 75 SHNGPDEAVDDKEDFDVPMVTQTSASAKIDYDKLIYRFGSSKIDKELINRIERATGQRPH 134
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120
DB 135 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 194
QY 121 LVIOQMTDEKYLKQDLTDOAYGDAVENKDIAGSPDINKTFISDLDYMGSSGFYKN 180
DB 195 LVIOQMTDEKYLKQDLTDOAYGDAVENKDIAGSPDINKTFISDLDYMGSSGFYKN 254
QY 181 VKIQRHTVFNQVKGIFGFTSDCIKISFPALQAPFSNSFPQIFRDRDIIQCLIPCA 240

```

```

Db 255 VKI0KVTENOVKGIFFGFTDSDICIGISFPALQAAFSFNSPFOIFHGADIQCLIPCA 314
QY 241 IDDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKOIKT 300
Db 315 IDDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKOIKT 374
QY 301 KANKHAFSGGRDTEEBHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGEL 360
Db 375 KANKHAFSGGRDTEEBHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGEL 434
QY 361 KKALEIYLOPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401
Db 435 KKEIIDLVLQPLVAEHQARRKVTDEIVKEFMTPRQLCFHQ 475

RESULT 4
SYM_MOUSE STANDARD; PRT; 481 AA.
ID SYM_MOUSE
AC P32521;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
DE WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA (Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: The short isoform is widely expressed, the
CC long form is found only in embryonic stem cells.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X69656; CA49347.1; -
DR EMBL; X69657; CA49348.1; -
DR PIR; S31461; S31461.
DR PIR; S31462; S31462.
DR MGD; MGI:104630; WARS.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PRO1039; TRNASYNTHTRP.
DR TIGRPFAMS; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA tRNA LIGASE_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
DR Anticodon-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.

```

```

FT SITE 353 357 "KMSKS" REGION.
FT VARSPLIC 476 481 MISSING (IN ISOFORM 2).
SQ SEQUENCE 481 AA; 54282 MW; B05M452C08074F52 CR664;
Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 7, 4e-152;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;
QY 2 NHCPDTEAEEDPVDVMTVQTSKAKIDYDKLIVRGSSKIDKELNRIERATGQRPPIHF 61
Db 76 NCDSATKASEDFVDPWTVRTSSAKIDYDKLIVGSSKIDKELNRIERATGQRPPIHF 135
QY 62 LRGGIFPSHRDMQVDAENKKPFYLYTGRGSSPAMVGHILPFIPTKMLQDVENVVL 121
Db 136 LRGGIFPSHRDMQVDAENKKPFYLYTGRGSSPAMVGHILVPIPTKMLQDVENVVL 195
QY 122 VIQMTDEKRYLWKDLTLQDAGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFEYKAV 181
Db 196 VIQMSDDEKRYLWKDLTLQDAGSYVENAKDIIACGFDINKTFIFSDLEVMGSGPGFEYKAV 255
QY 182 VKI0KVTENOVKGIFFGFTDSDICIGISFPALQAAFSFNSPFOIFHGADIQCLIPCAI 241
Db 256 VKI0KVTENOVKGIFFGFTDSDICIGISFPALQAAFSFNSPFOIFHGADIQCLIPCAI 315
QY 242 DDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKOIKTK 301
Db 316 DDDPYFRMTRDVAPRIGYKPKPALHSTFPALQAGATKMSADPNSSIFLTDTAKOIKTK 375
QY 302 VNKHAFSGGRDTEEBHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELK 361
Db 376 VNKHAFSGGRDTEEBHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELK 435
QY 362 KKALEIYLOPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401
Db 436 KKEIIDLVLQPLVAEHQARRKVTDEIVKEFMTPRQLCFHQ 475

RESULT 5
SYM_SCHRO STANDARD; PRT; 395 AA.
ID SYM_SCHRO
AC Q09692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TrpRS).
DE GN SPAC2F7.13C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RC MEDLINE=21648401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor A., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Welter J., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinbach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

```



RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revelata J.L., Moreno S., Armstrong J., Forburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of *Schizosaccharomyces pombe*,"  
 RL Nature 415:871-880(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diposphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Z50142; CA90500.1; -.  
 DR InterPro: IPR002306; Trp tRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs: TIGR00233; trpS; 1.  
 DR PROSITE: PS00178; AA tRNA\_LIGASE\_I; 1.  
 KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;  
 KW Ligase; Atp-binding.  
 FT SITE 91 100 "HIGH" REGION.  
 FT SITE 275 279 "KMSKS" REGION.  
 SQ SEQUENCE 395 AA; 44910 MW; E656AEBB76C5F9F9 CRC64;  
 Query Match 57.2%; Score 1210; DB 1; Length 395;  
 Best Local Similarity 59.6%; Pred. No. 4.2e-92;  
 Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;  
 QY 11 EEDFDPVPTVQTS---SAKGIDYDKLYRFGSSKIDKELINIERATGQRPHHFRGCI 66  
 DB 4 EEOIVTPMVGKSIDVGEKGIYERLIYQFGTRKITPQLERFEKLTGKKPHLLRGA 63  
 QY 67 FFSHDMNOVLDAVENKKPFYLYTGSGPSSEAMHGHLLPFIETKMLQOVFVPLVQMT 126  
 DB 64 FFSHDPDMITLIRYEEKKPFYLYTGSGPSDSHGLHGMIPFVCKMLQOVFVPLVQLT 123  
 QY 127 DDEKYLMD-LTLDDQAVGDAVENAKDIIACGPDINKTFIFSLDLYMGSSGFYKVVVKIQ 185  
 DB 124 DDEKFLFKQGVSLDQGRFARENAKDIIAVGPDPKTFIFMSTVYG--GAFYQVNVRIA 181  
 QY 186 KAVTNQVKGIFGFTSDICIGKISPAIDAAPSFNSFPQIRDRDIDLOICLPALDQDP 245  
 DB 182 KCIITNQSACGPFSDSDIGIKIHFAISIQAPSFSSPFHINCAIDICPLPCALDQDP 241  
 QY 246 YRMTRDVAPRIGYKPPALHSPFPALGAOTKMSASPNSSIFLTDFAKOIKTKVNG 305  
 DB 242 YRLTRDSVGRLLKFKKPPALHSPFPALQPOSKMSASOSKSIFFMTDIPNKIKKKNIN 301  
 QY 306 AFSGRDITIEHRQFQNCVAVSFMYLTFPLEDDDKLQIRKDYTSGMALTELKAL 365  
 DB 302 AFSGGATIEIRHEKGNPDVAVYLSFLDDDDKQLQVNTYVAGTSLGEMKGE 361  
 QY 366 EYLOPLIAEHQARKKAYDEIYKEFMT-PRKLSF 398  
 DB 362 KLLQFVSDFOAARSKYDEATLDMFMDGSRKLEW 395

DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--  
 DE tRNA ligase) (TrpRS).  
 GN WRS1 OR YOL097C OR HRE432.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_Taxid=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96076631; PubMed=7502582;  
 RA Vandenbol M., Durand P., Portetelle D., Hilger F.;  
 RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV  
 RT including the Tyl-H3 retrotransposon, the sufl1(+) frameshift  
 RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a  
 RT delta element.";  
 RL Yeast 11:1069-1075(1995).  
 RN [2]  
 RP FUNCTION.  
 RX MEDLINE=97197969; PubMed=9046085;  
 RA John T.R., Ghosh M., Johnson J.D.;  
 RT "Identification and expression of the *Saccharomyces cerevisiae*  
 RT cytoplasmic tryptophanyl-tRNA synthetase gene.";  
 RL Yeast 13:37-41(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diposphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Z48149; CA98164.1; -.  
 DR EMBL: Z74839; CA99110.1; -.  
 DR SGD: S0005457; WRS1.  
 DR InterPro: IPR002306; Trp tRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs: TIGR00233; trpS; 1.  
 DR PROSITE: PS00178; AA tRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; Atp-binding.  
 FT SITE 111 120 "HIGH" REGION.  
 FT SITE 295 299 "KMSKS" REGION.  
 SQ SEQUENCE 432 AA; 49350 MW; C408F169737E9736 CRC64;  
 Query Match 55.0%; Score 1163; DB 1; Length 432;  
 Best Local Similarity 54.8%; Pred. No. 3.5e-88;  
 Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;  
 QY 7 ATEAEDFVDPWTV-----QTSSAKGIDYDKLYRFGSSKIDKELINIERATGQRPHH 61  
 DB 19 STDVGEQVTPDVGVGVDGQGRAGNIDYDKLIRKFGTRPVAEETLKPFKQVYGRPHH 78  
 QY 62 LRGIFFSHRDMNOVLDAVENKKPFYLYTGSGPSSEAMHGHLLPFIETKMLQOVFVPL 121  
 DB 79 LRKGLFSEBRDPTKLLDIEQKPFYLYTGSGPSDSHGLHGMIPFVCKMLQOVFVPL 138  
 QY 122 VIQMTDDEKYLMD-LTLDDQAVGDAVENAKDIIACGPDINKTFIFSDIDYMGSSGFYKN 180  
 DB 139 VIELTDEKFLFKKLTITNDVKNFARENAKDIIAVGPDPKTFIFSDIQVWG--GAFYET 196  
 QY 181 VKIQKAVTNQVKGIFGFTSDICIGKISPAIDAAPSFNSFPQIRDRDIDLOICLPICA 240  
 DB 197 VVRVSRQITGSAKAVFGFNDSDICIGKHFASIQATAFSSFPVVLGLPKTCLIPICA 256

```

QY 241 IDDDYFMRTRDVAAPRIGYKPPALLHSTFPALQGAQTMSASDPNSIFLTDPAKOIKT 300
DB 257 IDDDYFMRTRDVAADKLKTSKPLLSRFPALQGAQTMSASDPNSIFLTDPAKOIKT 316
QY 301 KVKKHAFFSGGRDPIIEHRQFGNCDDVVSFMYLTFPLEDDDKLEQIRKDYTGAMLTGEL 360
DB 317 KINKAFVSGGQVADLHRELGNPDVVAQYLSFPPDDVFLKECDYKXSGELLSGEM 376
QY 361 KKALIEVLOPLIAEHQARKKEVTDVYKEMTRPKL 396
DB 377 KKLCTIELOEFVAKAFQERRAQVDEETLDFKMPVPHKL 412

RESULT 7
SYM_SUIISO STANDARD; PRT; 381 AA.
ID SYM_SUIISO
AC 0976M1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (trpS).
GN TRPS OR ST0169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jinno K., Takahashi M.,
RA Sekine M., Baba S.-I., Akai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Oguchi A.,
RA Aoki K.-I., Maeda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Ohnita T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Rep. 8:123-140(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; APO00981; BAB65126.1; ALT INIT.
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; tRNA-synt_1b.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC TIGRfam; TIGR00233; trpS; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I, FALSE NEG.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SITE 82 90 "HIGH" REGION.
CC SITE 254 258 "KMSKS" REGION.
CC FT SITE 381 AA; 44718 MW; 109FSA56AFD7J159 CRC64;
CC SEQUENCE

Query Match 43.0%; Score 910.5; DB 1; Length 381;
Best Local Similarity 50.5%; Pred. No. 1,7e-67;
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;
QY 12 EBF-VDPWYQTSAGK-IDYDKLIVRFGSSKIDKELINIEATGCRPHPLRGRGFFS 69
DB 3 ODENVTPWEV-----KQKVDYDKLIVQFGTKITSEIKKIKSIINDELHVMRLRDVFFS 57

```

```

QY 70 HRDMOVLDAENKKPFYLYTGRSPSEAMHGHILPFIETKMLDQVENVPIQWOTDE 129
DB 58 HRDLDLVLDVQDQGGKFFLYTGRAPSL-GMHLGHLPIFTMLDQKFNVLNLIETTDE 116
QY 130 KILWK-DLTLDAAYGAVENAKDIIACGFDINKTEIFSDLDYMGSSGFGYKNVVKIQKHV 188
DB 117 KFMNBEYTLDDGTQWAVYNNIIDITAVGNPNKTFIFQOTEYI---RNNYPIAIKIARKL 173
QY 189 TRNQVKGIGFGFSDSCIGKISFPALQGAQFNSPNSPQIRDRDITDQCLPCALIDDPYR 248
DB 174 TFSBVAATGTLDTSSNIGIIMYPALQIAET-----MEKKR--RCLIPAGIDDPYR 223
QY 249 MTRDVAAPRIGYKPPALLHSTFPALQGAQTMSASDPNSIFLTDPAKOIKTKVNHAFS 308
DB 224 LQDRIAESLGYKRAQIHKFLPPLTGPBGKSSSQEPFAIYLTDDPKTVERKIMKVFAS 283
QY 309 GGRDPIIEHRQFGNCDDVVSFMYLTFPLE-DDDKLEQIRKDYTGAMLTGELKKALIEV 367
DB 284 GGQPTIELRKYKGNPDIDVVSFQWLYMFPEDPDNKKIIEEDYRSGALLTGELKQITIEK 343
QY 368 LQPLIAEHQARKKE 381
DB 344 LNDFLIEHRQKREE 357

RESULT 8
SYM_SUIISO STANDARD; PRT; 380 AA.
ID SYM_SUIISO
AC 0972X0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (trpS).
GN TRPS OR SS00452.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awaysz M.J., Chan-Weber C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erasmo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Thieriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AEO06677; AAK40778.1; ALT INIT.
CC InterPro; IPR002306; Trp_cRNA-synt_1b.
CC InterPro; IPR002305; tRNA-synt_1b.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PR01039; TRNASYNTHTRP.
CC TIGRfam; TIGR00233; trpS; 1.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I, FALSE NEG.
DB

```

KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 81 89 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 380 AA; 44691 MW; CF8344CP6383680 CRC64;

Query Match  
 Best Local Similarity 42.8%; Score 905; DB 1; Length 380;  
 Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;

QY 11 EEDPVDPMVQTSSAKG-IDYDKLIVRFGSSKIDKELINRIERATGQRPHPFLRGIFFS 69  
 DB 3 DEFTVPMVEGV---KGKVDYDKLIVQFGTQKTEELKRIKMLAGDL-HVMLRRVYFSS 56  
 QY 70 HRDMQVLDAYENKRPFLYLTGSGSSAMHVGHLIPFTFKLQDVFNVLVIQMTDDE 129  
 DB 57 HBDLVLVNDYKESKQFLYTGAPSL-GHIGHLIPFTFKLQKFPANLYIETDDE 115  
 QY 130 KYLMK-DLTLDQAYGAVENAKDIIACGPDINTFIISDLIDYMGSSGFYKVVKIQKAV 188  
 DB 116 KVMRNEPFLDQSRSAVYNDIIVGNPDKTIFQDTEYI---RMVPTVYIAKTL 172  
 QY 189 TENVKGIQFTSDSCIGKISPPAIOAPSFNSPQIFRDRDIOCLIPCAIDQDPYR 248  
 DB 173 TSEVATFGLDASSNIGLIFYPALQIAPF-----MEKK---RCLIPAGIDQDPYR 222  
 QY 249 MTRDVAIRIGYPPRALLHSTFPALQAGQTKMSASDPNSSIFLTDAKQIKTVNKHAFS 308  
 DB 223 LQRDIAESIGYKAAQIHSKFPPLGPGKSSSNPEATYIVDPKVERIKMYAFS 282  
 QY 309 GGRDTEIHRQPGNCVDVVSFMYLTFPLEDDD-KLEQIRKDYTSAMTGLKALIEV 367  
 DB 283 GGQPTIELHRKYGNEIDVPPQMLYFFBEDNRIKEIEEYRSKMLTGLKQILIDK 342  
 QY 368 LQPLIAEHQARRKVTDEIVKEFTPKLS 397  
 DB 343 LNNFLEHR-RRREAKELVHVKYDGKLA 371

RESULT 9  
 ID\_SWM\_PYRAB STANDARD; PRT; 385 AA.  
 AC Q9UY11;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TPRS).  
 GN TRPS OR PAB111.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_Taxid=29292;  
 RN NCBI\_Taxid=29292;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GES / Orsay;  
 RA Hellig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +  
 CC -1- diphosphate + L-tryptophanyl-tRNA (Trp).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/annouce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

DR EMBL; AJ248288; CAB50601.1; -  
 DR InterPro; IPR002306; Trp-tRNA-synt\_1b.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTPR.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 82 90 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 385 AA; 45100 MW; 4C29D0141976B12 CRC64;

Query Match  
 Best Local Similarity 37.9%; Score 803; DB 1; Length 385;  
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 12 EDP-VDPMVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHPFLRGIFFS 70  
 DB 3 EDFKTPMVEGV---VDYNKLIHRGTSPLTELEKTALETSELPFRFRKFFPSH 58  
 QY 71 RDMQVLDAYENKRPFLYLTGSGSSAMHVGHLIPFTFKLQDVFNVLVIQMTDDE 130  
 DB 59 RYDKVLQDYEGRGFFLYTGRGSG-PWHIGHIIPFATKWLQKFGVNLVIQITDDEK 117  
 QY 131 YLMK-DLTLDQAYGAVENAKDIIACGPDINTFIISDLIDYMGSSGFYKVVKIQKAV 189  
 DB 118 FLFKNLTFFEDTKHAYENIIDIIVGPDPTFTFQNSRF---TKIYEMAPIAKKIN 173  
 QY 190 FQVQKGIQFTSDSCIGKISPPAIOAPSFNSPQIFRDRDIOCLIPCAIDQDPYR 249  
 DB 174 FSMARAVGFTQSKIGMIFFPALQIAPF-----PEKR---RCLIPALIDQDPYR 223  
 QY 250 TRDVAIRIGYPPRALLHSTFPALQAGQTKMSASDPNSSIFLTDAKQIKTVNKHAFS 309  
 DB 224 QRDFAESIGYKAAIHSKFPVSLTSLGKMSASKPEATYIVLTDSPEDVEKVVKAFALG 283  
 QY 310 GGRDTEIHRQPGNCVDVVSFMYLTFPLEDDD-KLEQIRKDY---TSGAMTGLKALIE 366  
 DB 284 GRPTLKEQREKGEPEKCVFWMLEIFFEEDK-KLMERYACKNGELTGCCKRYLS 341  
 QY 367 VLQPLIAEHQARRKVTDEIVK 388  
 DB 342 KIQEPLEKQKRRKKAQKQIEK 363

RESULT 10  
 ID\_SWM\_PYRHO STANDARD; PRT; 386 AA.  
 AC O59584;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TPRS).  
 GN TRPS OR PH1921.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_Taxid=53953;  
 RN NCBI\_Taxid=53953;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki U., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA (Trp).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in, no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP000007; BA31046.1; ALT. INIT.  
 DR InterPro: IPR002306; Trp\_cRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA TRNA\_LIGASE\_I; 1.  
 KM Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;  
 KM Complete proteome.  
 FT SITE 82 90 "HIGH" REGION.  
 FT SITE 253 257 "KMSKS" REGION.  
 SQ SEQUENCE 386 AA; 45305 MW; 9E3C392F4028B2D0 CRC64;  
 Query Match 37.6%; Score 796; DB 1; Length 386;  
 Best Local Similarity 44.0%; Pred. No. 4.5e-58;  
 Matches 172; Conservative 66; Mismatches 125; Indels 26; Gaps 9;  
 QY 11 EEDFVDPWTVQTSAGIDYDYLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSH 70  
 DB 3 EEFKVTPEWEVGV---VDYDKLIKHFQTSPLTDLKTAELTKSELPIFFRRKFFFSH 58  
 QY 71 RDMNQVLDAVENKKPFYLITGKGSSSEAMVGHILPIFTKMLQDVENVVLQMTDDEK 130  
 DB 59 RVDVLLIKDVEBGGFPLYTGKPGSG-PMHIGHILPFPATKMLQEKGVNLYIQITDDEK 117  
 QY 131 YLWMD-LITLQAVGDAVENAKDIIACGPDINKTIFPSDLIDYMGSSGFYKVVVIOKHT 189  
 DB 118 FLTKENLTFPDTKMAVDNLIDIIAVGFPDPKTIFQNSEF---TIYEMALPIAKIN 173  
 QY 190 FNOVKIGFETDSDICIGISFPALQAPSFNSFPQIFRDRTDIOCLIPCAIDODPYFRM 249  
 DB 174 FSNAAKAVFGFTGSKIGMIFFPALQIAPTF-----FEKK---RCLIPALIDODPYWRL 223  
 QY 250 TRVVAIRIGYPPKALHSTFPALQAGQTKMSASDPNSIFLTDIAKQITKYNKKAFFSG 309  
 DB 224 QRPFAEELGYKTAALHSEKVPSTLSGKMSASKEPETAIIYLTDSPEDEVKAKWKFLLTG 283  
 QY 310 GRDTIEHRQFGNGCDVDVSEFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 366  
 DB 284 GRDTLEHQRKSGEPKCVFPMKLETFEEDDK-KLKERIYACNGELTGCCKRILIS 341  
 QY 367 VLOPLIAHQARRKEVTDEIVKEFMTPRKLS 397  
 DB 342 KIQEFLKEHQRRRK-AEKLVEKFKYTKGKA 371  
 RESULT 11  
 ID SYM PYRAE STANDARD; PRT; 375 AA.  
 AC 082TUS;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)  
 DE (TrpRS).  
 OS TRPS OR PAE3091.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 NCBI TaxID=13773;

RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX PubMed=11792863;  
 RA Flitz-Gibbon S.T., Lader H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 RT aerophilum";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA (Trp).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE009911; AAL6464.1; -  
 DR InterPro: IPR002306; Trp\_cRNA-synt\_1b.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR TIGRFAMs: TIGR00233; trps; 1.  
 DR PROSITE: PS00178; AA TRNA\_LIGASE\_I; 1.  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; ligase; ATP-binding;  
 KM Complete proteome.  
 FT SITE 81 89 "HIGH" REGION.  
 FT SITE 258 262 "KMSKS" REGION.  
 SQ SEQUENCE 375 AA; 43178 MW; 3DDBF85DA680F16 CRC64;  
 Query Match 34.7%; Score 735; DB 1; Length 375;  
 Best Local Similarity 41.6%; Pred. No. 4.5e-53;  
 Matches 162; Conservative 79; Mismatches 130; Indels 18; Gaps 11;  
 QY 11 EEDFVDPWTVQTSAGIDYDYLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFS 69  
 DB 2 EEFKVTPEWEVGR---VDYDKLIKHFQAKPLTKDEVALLEKYAGE-VHPLIRRFVFA 56  
 QY 70 HRDMNQVLDAVENKKPFYLITGKGSSSEAMVGHILPIFTKMLQDVENVVLQMTDDE 129  
 DB 57 HRDFDPMKMGHGRPMALYTGKPGSG-PVHIGHWPMWILKWFSDKFGLEVYFQITDDE 115  
 QY 130 K-YLWMDLITLQAVGDAVENAKDIIACGPDINKTIFPSDLIDYMGSSGFYKVVVIOK 188  
 DB 116 KFYDDEMKLEETWNAVENALDVIALGSPERLHILDTKDI--KPLYPPIAVRVAKKL 172  
 QY 189 TFNOVKIGFETDSDICIGISFPALQAPSFNSFPQIFRDRTDIOCLIPCAIDODPYFR 248  
 DB 173 TMTVVAIRIGFTDSTNIGLIFPSLOIAVAFLPT--ELKREARPV--LIPCAIDODPYFR 228  
 QY 249 MTRDVAIRIGYPPKALHSTFPALQAGQTKMSASDPNSIFLTDIAKQITKYNKKAFFSG 308  
 DB 229 LAIDIADALGYPPSTLYSKFIALTG-ESKMSASNPDAIYLTDEKTVRRKV-WNAFT 286  
 QY 309 GGRDTIEHRQFGNGCDVDVSEFMYLTFPLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 368  
 DB 287 GGRPTAEQRKTYGNGNEVCVPFYHMLFPDDDSVKEIRQDCSGALLCGECKLKHEKI 346  
 QY 369 QPLIAHQARRKEVTDEIVKEFMTPRKLS 397  
 DB 347 TKFLKEHRRERKARSK-VDEYRLSVKLS 374  
 RESULT 12  
 ID SYM HALNI STANDARD; PRT; 380 AA.  
 AC 09HN66;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
GN (TrpRS)  
OC Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_Taxid=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Shroog J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weli R., Goo Y.A.,  
RA Lethauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddock D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isebaugher T.A., Peck R.F., Pohlschroder J., Spindich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,  
RA "Genome sequence of Halobacterium species NRC-1."  
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
RL -1- CARBOLYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC -1- diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: AE005109; AAG20355.1; -  
DR InterPro: IPR002306; Trp\_tRNA-synt\_1b.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRANSSYNTHTRP.  
DR TIGRFAMs: TIGR00233; ttps; 1.  
DR PROSITE: PS00178; AA TRNA LIGASE I; 1.  
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
KM Complete proteome.  
FT SITE 74 / 82 "HIGH" REGION.  
FT SITE 249 / 253 "KMSK" REGION.  
SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;  
Query Match 25.3%; Score 534.5; DB 1; Length 380;  
Best Local Similarity 35.4%; Pred. No. 1.4e-36;  
Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;  
QY 10 AEDFVDPWTQTSAGKIDYDKLIVRFSSKIDKELINRIRATGQRPHEFLRGIIPS 69  
DB 3 AGDNVTPYAVASDD---LDYEKTLARFGADELTDQGRAPF-----DHPVINGLFAVA 53  
QY 70 HDNMQVLDAYENKPEFYLYTGRGSEEMRNGHLIPFTKMLQDVENVPLVIQMTDE 129  
DB 54 GBDVDFLTAAGSOS---IVTGVGPGG-PMHGHANVVFYFARLQDFEFARYVPLSDSE 108  
QY 130 KYLMKDLTLDAQVDAVE--NAKDIACGFDINKTFIF---SDLDVW-GMSGSGFYKVVVI 184  
DB 109 KYMFDQYPAET-GDVLKRNADLAVGDDPELTRIVUTRDADVILPLATPAAGV--- 164  
QY 185 QKHTVFNOYKGIFFGFDSDICGKISFPALQAAPEFSNSFPQIRDRDTIOCTIPCAID 244  
DB 165 -RHATLQNYG-----EPDNVGAQAFYPAVQTAHL---LPQLVHG--EHETVPLIADDD 213  
QY 245 PYFRMTRDVAIRIGV--KPAHSHFFPALGCAQCKMSADPNSGIFLTDRAKOIKTV 302  
DB 214 PYFRSRDVAARARFVGRGALMQLPGLSLAG-PGRMSS--AGVISIRLTDSPDVTVREV 271  
QY 303 NGHAGSGGDTIEHRQFGANGCDVDVSFMVLTFFLEDD-KLEQIRKDYTSGLMGLSEL 361

DB 272 RHATYTGASVSEHRAGGVPAEDVPFOYLSAFEPDPAELARLEREYRAGDLSGLK 331  
QY 362 KALIEVLOPLAEHQARKEVTDEIVKEPMTPRKLSFD 399  
DB 332 DLAADRITIEFLAAHQRRRAALGD--VTEALDAFRILTDD 367  
RESULT 13  
ID SYM METUA STANDARD; PRT; 370 AA.  
AC 058810;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, last annotation update)  
DT 16-OCT-2001 (Rel. 40, last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
GN (TrpRS).  
OS TRPS OR MJ1415.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_Taxid=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.V., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Keriavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weissbrock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Sadowman J.L., Nguyen D.,  
RA Uitterback T.R., Kelley J.M., Peterson J.D., Fahlman P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";  
RL Science 273:1058-1073 (1996).  
CC -1- CARBOLYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC -1- diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: U67582; AAB99425.1; -  
DR TIGR: M01415; -  
DR InterPro: IPR002306; Trp\_tRNA-synt\_1b.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRANSSYNTHTRP.  
DR TIGRFAMs: TIGR00233; ttps; 1.  
DR PROSITE: PS00178; AA TRNA LIGASE I; FALSE NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KM Complete proteome.  
FT SITE 75 / 83 "HIGH" REGION.  
FT SITE 255 / 259 "KMSK" REGION.  
SQ SEQUENCE 370 AA; 42660 MW; E6C71107C82B59D CRC64;  
Query Match 19.4%; Score 409.5; DB 1; Length 370;  
Best Local Similarity 30.9%; Pred. No. 2.6e-26;  
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
QY 17 PWTQTSAGKIDYDKLIVRFSSKIDKELINRIRATGQRPHEFLRGIIPS HDNMQV 76  
DB 8 PW--ETPAV--IDYKTKMEQGVKPIVDVLDLKEE-----HFRFRNIILIGHRDPERI 57

```

QY 77 LDAYENKKPFLYLYGRGSPSEAMHVGHLPIFIKMLQ---DVFNVLVIOQMTDDEKYLK 133
DB 58 VDAIKNNKEBAVVSQMMPSGK-MHFGKMWVDLLKFKQXTNINIP-----ADLEAYMA 112
QY 134 KDITLDQAVGDAV-ENAKDIIACGFPDINKTFTISDDLDMGMSGGFYKNVYK-OKHVTEN 191
DB 113 RNNSPFTYKELALNEVNTYIALGLDPEKINVLQSKYQKV-----KDLILISKTNMS 167
QY 192 QVAGIGFTDSDICIGKISFPALQAPSFNSPQIFRDR--DIQCLIPCAIDODPFYRM 249
DB 168 EMATIGGFKETNIGHVAFIVQVADL---HPOLDENLSPEKPPVVPVIGIDDPHIL 224
QY 250 TRDVAER---IGYKPALHSTFEPALQAGTQMSASDPNSSIFLDTAKQIKTKYKMA 306
DB 225 TRDIARAKFKFPIPSSTYHRFMWTLGG--KMSSEKPEFALFLTDDEKTVKKIIFS-A 281
QY 307 FSGGRTIEHRQFG---NCDVDSFMYLTFFLEDDKLEQIRKDYTSAMLTGELKKA 363
DB 282 KTGREGTLEHKKYGGVPECVVYELFLY--HLILDKELAELIYOKCRSGELTGCKCKM 339
QY 364 LIEVLQPLAEHQARRKEVTDIELK 388
DB 340 AVERVEFLKDLKEKQAKELAVK 364

RESULT 14
SYM METTH STANDARD; PRT; 364 AA.
AC 026352;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpS)
GN TRPS OR MTH251.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pochiar B., Qiu D.,
RA Spadefora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patelwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genome.";
RT J. Bacteriol. 179:7135-7155(1997).
RL J. CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Tnp) = AMP +
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE00812; AAB84757.1;
CC InterPro; IPR002306; Tnp tRNA-synt_1b.
CC InterPro; IPR002305; tRNA-synt_1b.
CC InterPro; IPR001412; tRNA-synt_1.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PR01039; TRNASYNTHTRP.
CC TIGRfam; TIGR00233; trps; 1.

```

```

DR PROSITE; PS00178; AA tRNA LIGASE I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 70 "HIGH" REGION.
FT SITE 251 "KMSKS" REGION.
SQ SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;
Query Match 18.8%; Score 397.5; DB 1; Length 364;
Best local similarity 27.9%; Pred. No. 2.5e-25;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 15 VDPWVQTSANGIDYDKLIVRFGSSKIDKELINRERATGQRPHHFLRGIFFSRDMM 74
DB 2 IDPW-----GSAR-LEVDLIENFGVRF-SEVLDEV-----DEPSYLMRGIIIFGRDVE 50
QY 75 QVLDAYENKKPFLYLYGRGSPSEAMHVGHLPIFIKMLQDVFNVLVIOQMTDDEKYLK 134
DB 51 RIIISAKKGEDPAVVTGMMPSGR-MHFGKMWVDLLKFKQXTNINIP-----ADLEAYMA 108
QY 135 DLITDQAVGDAV-ENAKDIIACGFPDINK-----TFISDDLDMGMSGGFYKNVYK 183
DB 109 GVDFEDSRRIALIEVYIAGYIALGLDEKONHIVYLOSENLWEDLAVY----- 156
QY 184 IQGHTFQVKGIFGFTSDICIGKISFPALQAPSFNSPQIFRDRITQCLIPCAIQ 243
DB 157 LAGKVNFEELRAIYGFSTSMAMWYAFIIQVSDILHPOLDLEGPR---PIVVPGPQ 213
QY 244 DPEFMTTRVAPRI---GYKPALHSTFEPALQAGTQMSASDPNSSIFLDTAKQIK 299
DB 214 DPHIRLRIIDIAFRDRYFIIIPSSYHFMGLTGG--KMSNRPKSAIFLSDTPEEAE 271
QY 300 TKVKNKAFSGDRTIEHRQFGNCDVDSFMYLTFFLE-DDDKLEQIRKDYTSAMLTG 358
DB 272 AKI-RNAKTGGRFTLEKEGRELGVPECCIIYELLHYMGSDSRLEIYESCRNGTLMCG 330
QY 359 ELKALIEVLQPLAEHQARRKE 381
DB 331 ECKNNTAEIRIRKFEELSVYRK 353

RESULT 15
SYM ENCCU STANDARD; PRT; 134 AA.
AC 096771;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpS) (Fragment).
GN Encephalitozoon cuniculi.
OS Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277663; PubMed=9615449;
RA Peyretalade B., Brousseau V., Peyret P., Metenier G., Gouy M.,
RA Vivares C.P.;
RT "Microsporidia, mitochondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA (Tnp) = AMP +
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:28:00 ; Search time 15.5539 Seconds  
(without alignments)  
2478.466 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNHGDPATBAEDFVDFMTV.....VTDEIVKEFMTPRKLSFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2101	99.3	471	1 A41706	tryptophan-tRNA 11
2	2028.5	95.9	475	1 YWBO	tryptophan-tRNA 11
3	1946.5	92.0	475	1 YWRBPR	tryptophan-tRNA 11
4	1938	91.6	481	2 S50053	tryptophan-tRNA 11
5	1210	57.2	395	2 S58157	hypothetical prote
6	1163	55.0	432	2 S51901	tryptophan-tRNA 11
7	907	42.9	386	2 C90190	tryptophan-tRNA 11
8	803	37.9	385	2 C75020	tryptophan-tRNA 11
9	626.5	29.6	301	2 G71206	tryptophan-tRNA 11
10	534.5	25.3	380	2 G84373	tryptophan-tRNA 11
11	409.5	19.4	370	2 F64476	tryptophan-tRNA 11
12	397.5	18.8	364	2 E69131	tryptophan-tRNA 11
13	386	18.2	134	2 T43806	tryptophan-tRNA 11
14	370.5	17.5	420	2 E69461	tryptophan-tRNA 11
15	353	16.7	374	2 D72477	probable tryptophan
16	269.5	12.7	513	2 F84371	tryptophan-tRNA 11
17	192	9.1	323	2 H69346	tyrosyl-tRNA synth
18	188	8.9	364	2 E72512	probable tyrosyl-t
19	185.5	8.8	341	2 D95360	tryptophan-tRNA 11
20	185.5	8.8	341	2 G98125	tryptophan-tRNA 11
21	174.5	8.2	341	2 B86633	tryptophan-tRNA 11
22	172	8.1	394	2 A45999	tyrosine-tRNA 11ga
23	170.5	8.1	366	2 S75410	tyrosine-tRNA 11ga
24	169	8.0	395	2 H70385	tryptophan-tRNA 11
25	167	7.9	351	2 E75438	tryptophan-tRNA 11
26	166.5	7.9	346	2 B71496	tryptophan-tRNA 11
27	164.5	7.8	895	2 A86410	protein F3M18.22
28	163.5	7.7	353	2 E70100	tryptophan-tRNA 11
29	163	7.7	460	2 C84750	probable tyrosyl-t

30	159.5	7.5	337	2 F71300	tryptophan-tRNA 11
31	158.5	7.5	375	2 B75072	tyrosyl-tRNA synth
32	158	7.5	346	2 C81654	tryptophan-tRNA 11
33	154.5	7.3	102	2 T44994	probable tryptophan
34	152	7.2	344	2 H86590	tryptophan-tRNA 11
35	152	7.2	344	2 C72034	tryptophan-tRNA 11
36	149.5	7.1	327	2 C84374	tyrosyl-tRNA synth
37	149.5	7.1	408	2 T03741	probable tyrosine-
38	146	6.9	339	2 E64676	tryptophan-tRNA 11
39	146	6.9	375	2 F71093	tryptophan-tRNA 11
40	145	6.9	337	2 A11066	tyrosine-tRNA 11ga
41	138	6.5	319	2 H69102	tryptophan-tRNA 11
42	135.5	6.4	328	2 C72370	tyrosine-tRNA 11ga
43	135	6.4	365	2 E82052	tryptophan-tRNA 11
44	132.5	6.3	343	2 E73024	tryptophan-tRNA 11
45	132	6.2	328	1 YWBSF	tryptophan-tRNA 11

#### ALIGNMENTS

RESULT 1  
A41706  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homo  
C:Species: Homo sapiens (man)  
C>Date: 19-May-2000 #sequence revision 19-May-2000 #text change 03-Jun-2002  
C/Accession: A41633; A41706; S19246; UN0676; JH0533; S26287  
R:Flecker, J.; Rasmussen, H.H.; Juellesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A>Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma  
A:Reference number: A41633; MUID:92107982; PMID:1763065  
A/Accession: A41633  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-471 <RUB>  
A/Cross-references: GB:M77804; NID:G184656; PIDN:AAA67324.1; PID:G184657  
R:Buwit, U.; Flohr, T.; Boettger, E.C.  
EMBO J. 11, 489-496, 1992  
A>Title: Molecular cloning and characterization of an interferon induced human cDNA with  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A/Accession: A41706  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-471 <RUB>  
A/Cross-references: GB:M77804; NID:G184656; PIDN:AAA67324.1; PID:G184657  
R:Buwit, U.; Flohr, T.; Boettger, E.C.  
EMBO J. 11, 489-496, 1992  
A>Title: Molecular cloning and characterization of an interferon induced human cDNA with  
A:Reference number: S19246; MUID:92164636; PMID:1537332  
A/Accession: S19246  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-423, 'R', 425-471 <BUW>  
A/Cross-references: EMBL:X62570; NID:G32708; PIDN:CAA44450.1; PID:G32709  
A/Note: 213-ser and 214-Tyr were also found  
R:Frolova, L.Y.; Grigorjeva, A.Y.; Sudomoina, M.A.; Kiselev, L.L.  
Gene 128, 237-245, 1993  
A>Title: The human gene encoding tryptophan-tRNA synthetase: Interferon-response element  
A:Reference number: JN0676; MUID:9329292; PMID:7685728  
A/Accession: JN0676  
A/Molecule type: DNA  
A/Residues: 1-141, 182-471 <FROL>  
A/Cross-references: GB:X67918; GB:S62837; NID:G37968; GB:X67919; NID:G37969; GB:X67920; F  
4; NID:G37974; GB:X67925; GB:S62855; NID:G37975; GB:X67926; GB:S62856; NID:G37976; GB:X6  
A/Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residue 34  
R:Frolova, L.Y.; Sudomoina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kiselev, L.L.  
Gene 109, 291-296, 1991  
A>Title: Cloning and nucleotide sequence of the structural gene encoding for human trypt  
A:Reference number: JH0533; MUID:92112058; PMID:1765274  
A/Accession: JH0533  
A/Molecule type: mRNA

A/Residues: 1-212, 'GD', 215-471 <FRO2>  
 A/Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368  
 A/Experimental source: fibroblast  
 C/Genetics:  
 A/Gene: GDB:WARS; IEP53  
 A/Cross-references: GDB:119632; OMIM:191050  
 A/Map position: 14q23-14q31  
 A/Intons: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:19-64/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2101; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 2e-163;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPWTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60  
 DB 71 SNHGPDTEAEEDFVDPWTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 130  
 QY 61 FLRGGFFSHRDNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVP 120  
 DB 131 FLRGGFFSHRDNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVP 190  
 QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 180  
 DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 250  
 QY 181 VVKIQKHTFNQVKGIFGFTSDICIGISPPALQAPSFNSFPQIFRDRTDIOCLIPCA 240  
 DB 251 VVKIQKHTFNQVKGIFGFTSDICIGISPPALQAPSFNSFPQIFRDRTDIOCLIPCA 310  
 QY 241 IDDPYFRMTRDVAPRIGYKPKALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 300  
 DB 311 IDDPYFRMTRDVAPRIGYKPKALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 370  
 QY 301 KVNKAHFGSGRDTIEHRQFGNCNDVDSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 360  
 DB 371 KVNKAHFGSGRDTIEHRQFGNCNDVDSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 430  
 QY 361 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 401  
 DB 431 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 471

RESULT 2  
 YMO  
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - bovine  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Bos primigenius taurus (cattle)  
 C/Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text\_change 03-Jun-2002  
 C/Accession: A40279; JN0354; S10460; S14540  
 R/Garret, M.; Rajot, B.; Trezeguet, V.; Labouesse, J.; Merle, M.; Gandr, J.C.; Benedet  
 Biochemistry 30, 7809-7817, 1991  
 A/Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic  
 A/Reference number: A40279; MUID:91329348; PMID:1907847  
 A/Accession: A40279  
 A/Molecule type: mRNA  
 A/Residues: 1-475 <GAR>  
 A/Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799  
 A/Experimental source: pancreas  
 A/Note: The authors translated the codon CTG for residue 347 as Ala and CAG for residue  
 R/Zargartava, T.A.; Kovalova, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.  
 Biorg. Khim. 15, 1307-1311, 1989  
 A/Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca  
 A/Reference number: JN0354; MUID:90211408; PMID:2631684  
 A/Accession: JN0354  
 A/Molecule type: protein  
 A/Residues: 112-124;282-287,'N',288,'F',289-292,'Q',293-294,'IR',336-353;423-441,443-445  
 A/Experimental source: liver  
 A/Note: this paper is in Russian  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C/Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:24-69/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 95.9%; Score 2028.5; DB 1; Length 475;  
 Best Local Similarity 95.8%; Pred. No. 1.7e-157;  
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SNHGPDTEAEEDFVDPWTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60  
 DB 76 SGEGLATEADEDFVDPWTVQSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 135  
 QY 61 FLRGGFFSHRDNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVP 120  
 DB 136 FLRGGFFSHRDNQVLDAVENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVFNVP 195  
 QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGFYKN 180  
 DB 196 LVIQMTDEKYLKMDLTLDQAYGDAVENAKD-ITCCFDINKTFIFSDLDYMGSSGFYKN 254  
 QY 181 VVKIQKHTFNQVKGIFGFTSDICIGISPPALQAPSFNSFPQIFRDRTDIOCLIPCA 240  
 DB 255 VVKIQKHTFNQVKGIFGFTSDICIGISPPALQAPSFNSFPQIFRDRTDIOCLIPCA 314  
 QY 241 IDDPYFRMTRDVAPRIGYKPKALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 300  
 DB 315 IDDPYFRMTRDVAPRIGYKPKALHSTFPALQAGTQMSADPNSSIFLDTAKQIKT 374  
 QY 301 KVNKAHFGSGRDTIEHRQFGNCNDVDSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 360  
 DB 375 KVNKAHFGSGRDTIEHRQFGNCNDVDSFMYLTFLEDDDKLEQIRKDYTSAMLTGEL 434  
 QY 361 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 401  
 DB 435 KKAILEVLOPLIAEHQARRKEVTDEIVKEFMPRKLSFDFQ 475

RESULT 3  
 YMRBP  
 tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 30-Sep-1992 #sequence revision 13-Feb-1998 #text\_change 03-Jun-2002  
 C/Accession: A35904; S37396  
 R/Lee, C.C.; Craigden, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990  
 A/Title: Cloning and expression of a mammalian peptide chain release factor with sequence  
 A/Reference number: A35904; MUID:90239043; PMID:2185472  
 A/Accession: A35904  
 A/Molecule type: mRNA  
 A/Residues: 1-475 <LEB>  
 A/Cross-references: GB:M33460  
 R/Polova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegen, G.; McCaughan, K.K.;  
 EMBO J. 12, 4013-4019, 1993  
 A/Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase are  
 A/Reference number: S37396; MUID:94009008; PMID:8404867  
 A/Accession: S37396  
 A/Molecule type: mRNA  
 A/Residues: 166-177 <FRO>  
 C/Genetics:  
 A/Gene: WRS  
 C/Complex: homodimer [validated, MUID:94009008]  
 C/Function:  
 A/Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent format  
 A/Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain release  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protein bi  
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>  
 F:174-177/Region: ATP-binding motif (HXGH)

Query Match 92.0%; Score 1946.5; DB 1; Length 475;  
 Best Local Similarity 90.8%; Pred. No. 8.2e-151;  
 Matches 364; Conservative 21; Mismatches 15; Indels 1; Gaps 1;

QY 2 NHG-PDATEABEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELINRIBATGQRPBH 60  
 Db SHDDPEAVDDKEDFVDPWTVQTSAGKIDYDKLIVQSSKIDKELINRIBATGQRPBH 134  
 QY 61 FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPL 120  
 Db FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPL 134  
 QY 121 LVQMTDDEKYLMDLTLDOAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGYYKN 180  
 Db LVQMSDDEKYLMDLTLDOAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGYYKN 254  
 QY 181 VKIQKHVTVNQVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIQCCLIPCA 240  
 Db VKIQKHVTVNQVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIQCCLIPCA 314  
 QY 241 IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSIFLTDTAKQIKT 300  
 Db IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSIFLTDTAKQIKT 374  
 QY 301 KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYSSGAMLTGEL 360  
 Db KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYSSGAMLTGEL 434  
 QY 361 KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401  
 Db KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 475

## RESULT 4

550053  
 tryptophan-tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse

C/Species: Mus musculus (house mouse)  
 C/Date: 07-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 03-Jun-2002

C/Accession: S50053; S50052; I49391; S31461; S31462  
 R/Pajot, B.; Sargier, C.; Bonnet, J.; Garret, M.

J. Mol. Biol. 242, 599-603, 1994

A/Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthetase  
 A/Reference number: S50052; MUID:95018226; PMID:7932716

A/Accession: S50053

A/Molecule type: mRNA  
 A/Residues: 1-481 <PAJ>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

A/Genetics: LSP  
 A/Note: Intron position was determined by sequencing of genomic DNA

A/Accession: S50052

A/Molecule type: mRNA  
 A/Residues: 1-475 <PAJ>

A/Cross-references: EMBL:X69656; NID:G55435; PIDN:CAA49347.1; PID:G55436

A/Genetics: SSP  
 R. Kiesel, L.L.

Biochimie 75, 1027-1039, 1993

A/Title: Mammalian tryptophanyl-tRNA synthetases.

A/Reference number: I49391; MUID:94257729; PMID:7515282

A/Accession: I49391

A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA

A/Residues: 1-481 <RES>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

C/Genetics: <LSP>

A/Genetics: WRS

A/Introns: 475/2

A/Note: the list of introns may be incomplete; clone W13

C/Genetics: <SSP>

A/Genetics: WRS

C/Suprafamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C/Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis

F.1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted  
 F.1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted  
 F.1-3-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 91.6%; Score 1938; DB 2; Length 481;

Best Local Similarity 90.8%; Pred. No. 4, 1e-150;  
 Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 NHG-PDATEABEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELINRIBATGQRPBH 61  
 Db NCDSDATKASEDFVDPWTVQTSAGKIDYDKLIVQSSKIDKELINRIBATGQRPBH 135  
 QY 62 FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPL 121  
 Db FLRRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPL 135  
 QY 121 LVQMTDDEKYLMDLTLDOAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGYYKN 181  
 Db LVQMSDDEKYLMDLTLDOAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGYYKN 255  
 QY 181 VKIQKHVTVNQVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIQCCLIPCA 241  
 Db VKIQKHVTVNQVKGIFGFTSDICIGKISFPALQAPSPNSFPQIFRDRTDIQCCLIPCA 315  
 QY 241 IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSIFLTDTAKQIKT 301  
 Db IDDDPYFRMTRDVAPRIGYKRPALHSTFPALQAGATKMSADPNSIFLTDTAKQIKT 375  
 QY 301 KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYSSGAMLTGEL 361  
 Db KVKHAFSGGRDITIEHRQFGNCVDVSPMYLTFPLEDDDKLEQIRKDYSSGAMLTGEL 435  
 QY 361 KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 401  
 Db KKALIEVLQPLIAEHQARRKVTDEIVKEFMTPRKLSFDFQ 475

## RESULT 5

558157  
 hypothetical protein SPAC2F7.13c - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe  
 C/Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 10-Dec-1999

C/Accession: S58157; T38561

R/Gentles, S.; Church, C.M.  
 submitted to the EMBL Data Library, July 1995

A/Reference number: S58145

A/Accession: S58157

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-395 <GEN>

A/Cross-references: EMBL:Z50142; NID:G1052783; PIDN:CAA90500.1; PID:G1052796

R/Gentles, S.; Church, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, July 1995

A/Reference number: Z21799

A/Accession: T38561

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-395 <GE2>

A/Cross-references: EMBL:Z50142; PIDN:CAA90500.1; GSPDB:GN00066; SPDB:SPAC2F7.13c

A/Experimental source: strain 972h-; cosmid c2F7

C/Genetics: SPDB:SPAC2F7.13c

A/Map position: 1

C/Suprafamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 57.2%; Score 1210; DB 2; Length 395;  
 Best Local Similarity 59.6%; Pred. No. 7, 2e-91;  
 Matches 225; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 11 EEDFVDPWTVQTS---SAKGIDYDKLIVRGSSKIDKELINRIBATGQRPBHFLRRGI 66  
 Db EEOIVTPMVQKSIYDGEKGIYDVERLIVQFTRKITPRQLERPEKLTCKKHLLRRGA 63  
 QY 67 FFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPLVQMT 126  
 Db FFSHRDMNOVLDAENKKPFYLYTGRGSSAMVGHILPFIPTKMLQDVNVPLVQMT 123  
 QY 127 DDEKYLMDLTLDOAYGDAVENADIIACGPDINKTFIFSLDYMGMSSGYYKAVVVKIQ 185

```

Db 124 DDEKFLFKQGVSLDDCCRFAPENAKDIIAAGFDPKTFIFMNSYTV--GAFQNVRIA 181
QY 186 KHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCAIDDP 245
Db 182 KCTTANQSKACGFTDSDISGKIFHAFSIOAAPSSSFPHIFNGAKDIPCLIPCAIDDP 241
QY 246 YFMTDVAARIGPKPALLHSTFPFALQAQOTMSADPNSSIFLDTAKQIKTKNKH 305
Db 242 YFRLTRVSGRLFKPKALLHSTFPFALQAQOTMSADPNSSIFLDTAKQIKTKNKH 301
QY 306 AFGSGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGELKKAL 365
Db 302 AFGSGGATIEHRKGNPNVDVAQYLSFLDDELKQIKYNTKAGTISTGEMKECI 361
QY 366 EVLQPLIAEHQARRKEVTDEIVKEFMT-PRKLSF 398
Db 362 KTLQGFVSDFOAARSKVDEATLDMFMDGSRKLEW 395

```

## RESULT 6

551901  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - Yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein HR432; protein 00792; tryptophanyl-tRNA synthetase  
C/Species: *Saccharomyces cerevisiae*  
C/Date: 05-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 03-Jun-2002

C/Accession: S51901; S59177; S66793  
R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
submitted to the EMBL Data Library, January 1995

A/Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including and a Delta.  
A/Reference number: S51848

A/Accession: S51901  
A/Molecule type: DNA

A/Residues: 1-432 <DNA>

A/Cross-references: EMBL:Z48149; NID:9663234; PIDN:CAA88164.1; PID:9663256

R/Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
Yeast 11, 1069-1075, 1995

A/Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the a delta element.  
A/Reference number: S59156; MUID:96076631; PMID:7502582

A/Accession: S59177  
A/Reference number: S59156; MUID:96076631; PMID:7502582

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-432 <DNA>

A/Cross-references: EMBL:Z48149; NID:9663234; PIDN:CAA88164.1; PID:9663256

A/Note: the nucleotide/sequence was submitted to the EMBL Data Library, January 1995

R/Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.  
submitted to the Protein Sequence Database, July 1996

A/Reference number: S66791  
A/Accession: S66793

A/Molecule type: DNA

A/Residues: 1-432 <DNA>

A/Cross-references: EMBL:Z74839; NID:91419947; PIDN:CAA99110.1; PID:91419948; GSPDB:GN00

A/Experimental source: strain S288C

C/Genetics:  
A/Gene: SGD:WRS1, WRS1, MIPS:YOL097C  
A/Cross-references: SGD:S0005457

A/Map position: 15L  
C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C/Keywords: ligase  
F:117-120/Region: ATP-binding motif (HXHG)

Query Match 55.0%; Score 1163; DB 2; Length 432;

Best local similarity 54.8%; Pred. No. 5.6e-87;  
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

```

QY 7 ATEAEEDFVDPWVY-----QTSSAKGIDYDKLIVRFSSKIDKELINRIRATGQRPHHF 61
Db 19 STDVKEGVVTPWDEGGVDGGRQNIIDYDKLIVRFSSKIDKELINRIRATGQRPHHF 78
QY 62 LRGIIPSHSDMNQVLDAYENKKPFYLYTGKGPSSSEAMHVGHLPIFTKMLQDVFNVP 121

```

```

Db 79 LRKGLFESRDFTKLIDLYEQGKPFYLYTGKGPSSSDSMELGHMIPFVFTKMLQEVDPVL 138
QY 122 VIQMTDEKYLWK-DLTLDQAYGDAVENAKDIIAGCFDINKFTIFESDLDMGSSGFYKA 180
Db 139 VIELTDEKFLFGKRLTINDVKNFAENAKDIIAGCFDINKFTIFESDLDMG--GAFYET 196
QY 181 VVKIQHVTFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCA 240
Db 197 VAVSRQITGSRAKAVFGNRSDCIGKHFASIQITAPSSFPNVLGDPKTPCLIPCA 256
QY 241 IDDPYFMTDVAARIGPKPALLHSTFPFALQAQOTMSADPNSSIFLDTAKQIKT 300
Db 257 IDDPYFRCRDVADLTKTSKPALHSTFPFALQSTTMSADPNSSIFLDTAKQIKT 316
QY 301 KYNKAFSGGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGEL 360
Db 317 KINKVAFSGQVADLHRELGPNVDVAQYLSFLDDELKQIKYNTKAGTISTGEMKECI 376
QY 361 KKAILEVQPLIAEHQARRKEVTDEIVKEFMT-PRKLSF 396
Db 377 KKLCTETLQEFVAKAFQERRAQVDEETLDKFMVPHKL 412

```

## RESULT 7

C90190  
tryptophanyl-tRNA synthetase (trps) [imported] - *Sulfolobus solfataricus*

C/Species: *Sulfolobus solfataricus*  
C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001

C/Accession: C90190  
R/She, Q.; Shing, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Y

arrst, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.  
submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.  
A/Reference number: A99139

A/Accession: C90190  
A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1386 <NR>

A/Cross-references: GB:AE006641; NID:g13813608; PIDN:AAK40778.1; GSPDB:GN00155

C/Genetics:  
A/Gene: trps

C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 42.9%; Score 907; DB 2; Length 386;

Best local similarity 48.6%; Pred. No. 3.6e-66;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

```

QY 8 TEAEEDFVDPWVYQVTSAGK-IDYDKLIVRFSSKIDKELINRIRATGQRPHHFIRGI 66
Db 6 TMDDEFTVTPWEY-----KGYVDYDLIVQFGTKITBELKQRIKMLADL-HVTLRRNV 59
QY 67 FFSHRDMNQVLDAYENKKPFYLYTGKGPSSSEAMHVGHLPIFTKMLQDVFNVP 126
Db 60 FFSHRDLVDLVNDYKSGKFLYTGRAPEL-GMIGHILPIFTKMLQKFNANLYEIT 118
QY 127 DDEKYLWK-DLTLDQAYGDAVENAKDIIAGCFDINKFTIFESDLDMGSSGFYKYNV 185
Db 119 DDEKYNRNEFTLIDQRSWAVYDNIIDIVGFPDPTFIFQDEYI---RNMPITVKA 175
QY 186 KHTVFNQVKIGFPTSDCIGKISFPALQAAPSFNSFPQIFRDRDIOCLIPCAIDDP 245
Db 176 KLTTFSEVATRELDASSNIGLIFYPALQIAPF-----MFEKK---KCLIPAGIDDP 225
QY 246 YFMTDVAARIGPKPALLHSTFPFALQAQOTMSADPNSSIFLDTAKQIKTKNKH 305
Db 226 YMLRQDIASIGYKAAQIHSKFLPLTGPBGMSSNPDEPAIYLVDPKTYERKIMY 285
QY 306 AFGSGRDTIEHRQFGNCVDVVSFMYLTFLEDDDLKQIRKDYTSGAMLTGELKKAL 364
Db 286 AFGSGQPTIELHRRKYGNPIDVDFQWLYFPEEDNRIKEIEEYRSGKMLTGELKKAL 345
QY 365 IEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLS 397

```

Db 346 IDLNNFLEBHR-RRBEAKELVHFVXDKLA 377

## RESULT 8

C75020

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: C75020

R:anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: C75020

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 &lt;KAW&gt;

A:Cross-references: GB:AJ248288; GB:AL096836; NID:G5458960; PIDN:CAB50601.1; PID:G545911

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: trps; PAB1111

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 37.9%; Score 803; DB 2; Length 385;

Best Local Similarity 45.3%; Pred. No. 1,1e-57;

Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 12 EDF-VDPMVQTSASAGIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFSH 70

Db 3 EDFKVPMEVEGV---VDYNKLIHFHGTSPLEELTEKTAELTKSELPLFFRRKKEFFSH 58

QY 71 RDMNOVLDAVENKKPFLVLTGRGSSSEAMVHGLIPFTKMDVENVPLVQMDDEK 130

Db 59 RYDKVLQDIEBGRGFLVLTGRGSSG-PMHIGHILFFPAIKWQERFGVNLVQITDDK 117

QY 131 YLMKD-LTLDOAYGDAVENAKDIIACGPDINKTFIESDLDYMGSSGFYKVVKIQKHVT 189

Db 118 FLFKENLTEDTGMAYENILDIANGPDPTKFIQNSF---TKIYEMAIPIAKKIN 173

QY 190 FNOVKGIFGPTSDCTGKISFPALQAPSFNSFPOIFRDRITDIOCLIPCAIDDDPYFM 249

Db 174 FMAKAVFEFTQESKIGMIFFPALQAPF-----FEKR---RCLIPALIDDDPYMRL 223

QY 250 TSDVAPRIVYPRKALHSTFPFALQAGTMSADNSFILTDTAKQIKTKNKAFAFG 309

Db 224 QDDPASELIGYKPAALHSTFVPSLTSLSGMSASKETAIYLTDSPEDEKVKWKFPALTG 283

QY 310 GRDTIEHRQFGNCVDVSPFMTLTFELEDDELBOIRKDY---TSGAMLTELKKAALIE 366

Db 284 GRPLTEKQEKGEPEKCVFVKMLTFELEDK--KLKERYVACKNGELTGCCKRYLLS 341

QY 367 VLQPLAEHQARKKEVTDEIVK 388

Db 342 KIQEFLKEHQKRRKKAKQIEK 363

## RESULT 9

G71206

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 03-Jun-2002

C:Accession: G71206

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatake, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushiida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; NUID:98344137; PMID:9679194

A:Accession: G71206

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-301 &lt;KAW&gt;

A:Cross-references: GB:AP000007; NID:G3236134; PIDN:BA431046.1; PID:G3258363

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH1921

C:Superfamily: yeast tyrosine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 29.6%; Score 626.5; DB 2; Length 301;

Best Local Similarity 44.2%; Pred. No. 1.9e-43;

Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

QY 99 MHVGHILPIPTKMLQDVFNVPVLIQMTDEKTLMKD-LTLDOAYGDAVENAKDIIACG 157

Db 1 MHIGHIIFPAIKWQERFGVNLVQITDDKELFESNLTDPDTKMAVDNIIDITAVGF 60

QY 158 DINKTFIESDLDYMGSSGFYKVVKIQKHVTENQYKIGFTSDCTGKISFPALQAP 217

Db 61 DDKTPTFQNSF---TKIYEMAIPIAKKINSMAKAVGFTQSKIGMIFFPALQIAP 116

QY 218 SPNSNFPQIFRDRITDIOCLIPCAIDDDPYFRMTVDVAPRIGYPRKALHSTFPALQAG 277

Db 117 TF-----FEKR---RCLIPALIDDDPYMRQDPFASISLGYKTAALHSTFVPSLTSL 166

QY 278 TKMSADPNSSIFLDTAKQIKTKVKGAFSGGRDTIEHRQFGNCVDVSPFMTLTFPL 337

Db 167 GMSASKETAIYLTDSPEDEKVKWKFTLTGRFTLKQREKGEPEKCVFVKMLTIF 226

QY 338 EDDKLEQIRKDY---TSGAMLTELKKAALIEVLQPLAEHQARKKEVTDEIVKEFMTPR 394

Db 227 EEDDK--KLKERYVACKNGELTGCCKRYLLSKIQFLKEHQRRKK-AKLYEKRYTG 283

QY 395 KLS 397

Db 284 KLA 286

## RESULT 10

G84373

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: G84373

R:Ng, W.V.; Kennedy, S.P.; Mahatas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.

; Leichner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor

Jung, K.H.; Alam, M.; Fretz, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; NUID:20504483; PMID:11016950

A:Accession: G84373

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 &lt;STO&gt;

A:Cross-references: GB:AE004437; NID:G10581646; PIDN:AA20355.1; GSPDB:GN00138

C:Genetics:

A:Gene: trps2

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 25.3%; Score 534.5; DB 2; Length 380;

Best Local Similarity 35.4%; Pred. No. 8.3e-36;

Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

QY 10 AEDPVDPMTVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFS 69

Db 3 ADGNVTPYAVSD---LDYKELARFADDELTDQARFP-----DHPVNLGLFYA 53

QY 70 HDMNOVLDAVENKKPFLVLTGRGSSSEAMVHGLIPFTKMLQDVFNVPVLIQMTDE 129

Db 54 GRVDVDFLTAGEQS---ITVGVPSSG-PMHIGHILFFPAIKWQERFGVNLVQITDDK 108

QY 130 KYLMKDLTLDOAYGDAVE-NAKDIIACGPDINKTFI---SDLDYM-GMSGSGFYKVVKI 184

Db 109 KVMFQDQIPAEF-GDYLRANLRDLLAVGFDPDELTVIVDTDRADVLYPLATFADVD--- 164

QY 185 QKVTNQKIGIFGTFDSDICIGISPAIOAPSFSNSPFOIFRDRDIOCLIPCAIDDD 244  
 DB 165 -RHATQONVVG-----EPDVGQAFYPAVQTAHL--LQOLVHG--EHETLVPIADDD 213  
 QY 245 PYRMTDVAIPRIQYP--KEPALHSTFFPALOGAQTMSASDNPSSIFLTDTRAKQITKV 302  
 DB 214 PHRVSRDVAAKARYPYKGRGALLMQLPBLAG--PGKMSSS--AGVSIRLTSDPDYREKV 271  
 QY 303 NKGAIFSGRPTIEBHROFGNCVDVVSFWTLTFLEDD--KLEQIRKDYSGAMLTGELK 361  
 DB 272 RTAAHYGGRASVEBHRAKGVPAEDVPQYLSAFEPDDAEIARIERYRAGDLSGELK 331  
 QY 362 KALIEVLQPIAEHQARKEVTEIYKEFMTPRRLSPD 399  
 DB 332 DLAAADRIEFLAHQRRRALGD--VTEALDARLITDD 367

RESULT 11  
 F64476  
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Methanococcus jannaschii  
 C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 03-Jun-2002  
 C/Accession: F64476  
 R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Tron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hult, M.A.  
 Science 273, 1058-1073, 1996  
 A/Authors: Kahne, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C.  
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A/Reference number: A64300; MUID:96337999; PMID:868087  
 A/Accession: F64476  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-370 <BUL>  
 A/Cross-references: GB:U67582; GB:L77117; NID:g1592064; PIDN:AAB99425.1; PID:g1592065; T  
 C/Genetics:  
 A/Map position: FORJ375885-1376997  
 A/Start codon: GTG  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 19.4%; Score 409.5; DB 2; Length 370;  
 Best Local Similarity 30.9%; Pred. No. 1.3e-25;  
 Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

QY 17 PWTQTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDNNQV 76  
 DB 8 PW--ETPAV--IDYKTMEOFGVPIYDVLGDLKEE-----HHFRRNITILGHRDPERI 57  
 QY 77 LDAYENKKPFLYLTGRGSSSEAMVGHILPFIPTKMQ--DVFNVLVLOMTDDEKYLW 133  
 DB 58 VDAIKNNKEPAVVSQGMPSGR-MHFGKMWVLDLKFQKTYDININPI----ADLEAYMA 112  
 QY 134 KDLTLDAVGDVAV--ENAKDIIACGFDINKTFIFSDLDYMGSSGFYNNVKI--QKVTFN 191  
 DB 113 RNSSFETTKLALNEVITNYIALGLDPEKINVIQSKYQV----KDLALLISKRTNWS 167  
 QY 192 QVKGIFGTFDSDICIGISPAIOAPSFSNSPFOIFRDRD--DIQCLIPCAIDDDPYRNM 249  
 DB 168 EMKATVGFKGFTNIGHFAPIVQADIL--HPQDLNLSPEPKPVVAVPGIDDDPHIRL 224  
 QY 250 TRVAVP--IGYKPAALHSTFFPALOGAQTMSASDNPSSIFLTDTRAKQITKVAKHA 306  
 DB 225 TRDIANNAKEPKFTIPSSSTYHRFWTGLLG--KMSSSKPEPTALFLTDDEKTKKIKFS-A 261  
 QY 307 FSGGRDIEBHROFG--NCDVVSFWTLTFLEDDDKLEQIRKDYSGAMLTGELKKA 363  
 DB 282 KTGKRETLIEHKKYGVPECVVYELFLY--HLILDKEIAELIYKGRSGSELTCGKCKKM 339  
 QY 364 LIEVLOPIAEHQARKEVTEIYK 388

DB 340 AYERVVEFLKDLKEKREQAETIAVK 364

RESULT 12  
 E69131  
 tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta  
 N/Alternate names: tryptophanyl-tRNA synthetase  
 C/Species: Methanobacterium thermoautotrophicum  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Jun-2002  
 C/Accession: E69131  
 R/Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; I  
 Jiu, D.; Spadafora, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Utwani, N.;  
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi  
 A/Reference number: A69000; MUID:98037514; PMID:9371463  
 A/Accession: E69131  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-364 <MTH>  
 A/Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84757.1; PID:g2621301  
 C/Genetics:  
 A/Experimental source: strain Delta H  
 A/Gene: MTH251  
 A/Start codon: TTG  
 C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.8%; Score 397.5; DB 2; Length 364;  
 Best Local Similarity 27.9%; Pred. No. 1.2e-24;  
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 15 VDPWQTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDNN 74  
 DB 2 IDPW----GSAK-LEIYQDILNFGVPRP-SEVLDVY-----PEPSLMRGIIFGRDYE 50  
 QY 75 QVLDAYENKKPFLYLTGRGSSSEAMVGHILPFIPTKMLODVFNVLVLOMTDDEKYLW 134  
 DB 51 RIISAMKKGEDFAVVTGMMPSGR-MHIGKMIYDQLRW-YDRGAGRIPIPIDMEAYSR 108  
 QY 135 DLTLDAYGDVAVEN-AKDIACGFDINK-----TFISDLDYMGSSGFYNNVK 183  
 DB 109 GVPEDSRRIAEIHYIAGYIALGLDLEKDNHIVYLOSENLMWEDLAVV----- 156  
 QY 184 IOGHVFNQVKGIFGTFDSDICIGISPAIOAPSFSNSPFOIFRDRDIOCLIPCAIDDD 243  
 DB 157 LACKVFNELRALYGTGTSMAHMYAPILIOVSDILHPQDLDELGRP---PVIVPGPQD 213  
 QY 244 DPEFMTDVAIPRI--GYKPAALHSTFFPALOGAQTMSASDNPSSIFLTDTRAKQIT 299  
 DB 214 DPHIRLTRDIAARFRDRYGFILPSSSTYHRFMGLTGG--KMSSNRKSAIFLSDTPEEAB 271  
 QY 300 TKYNKAIFSGRPTIEBHROFGNCVDVVSFWTLTFLE--DDDKLEQIRKDYSGAMLTG 358  
 DB 272 AKI-RNAKTGGRTELEORELGQVPEECIIYETLLVHMGSSDSRLEIYESCRNLTGCG 330  
 QY 359 ELKALIEVLQPIAEHQARKE 381  
 DB 331 ECKANTAEIRKFFEEISVAREK 353

RESULT 13  
 T43806  
 tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)  
 C/Species: Encephalitozoon cuniculi  
 C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 03-Jun-2002  
 C/Accession: T43806  
 R/Peyretallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivaree, C.P.  
 Mol. Biol. Evol. 15, 683-689, 1998  
 A/Title: Microsporidia, mitochondrial proteins, possess a 70-kDa heat shock protein gene  
 A/Reference number: Z22693; MUID:98277683; PMID:9615449  
 A/Accession: T43806  
 A/Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA  
 A.Residues: 1-134 <PEV>  
 A.Cross-references: EMBL:AJ012470; PIDD:CAA10034.1  
 C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C.Keywords: ligase

Query Match 18.2%; Score 386; DB 2; Length 134;  
 Best Local Similarity 51.5%; Pred. No. 2.5e-24;  
 Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 12 EDFVDPWTQTSAAK---GIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIF 67  
 DB 3 EGRITPMDVEAVSTDEVPALIDYDKLINOFGCEKFNQALADRLKLSGKPAHFGRGIV 62  
 QY 68 FSHRDNNQVLDAVENKKPPLYLTGSGPSSSEAMHVGHLIFFTKMLQDVFNPLVYIQMD 127  
 DB 63 FMRDNNLLDDEIANNRPPLYLTGSGPSSKTMHIGHTIFPLCKYQMDAFKRLVLIQITD 122  
 QY 128 DEKYLWKDITLD 139  
 DB 123 DEKFLKMSWLE 134

#### RESULT 14

B69461  
 tryptophanyl-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C.Species: Archaeoglobus fulgidus  
 C.Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C.Accession: B69461

R.Klenk, H.P.; Claydon, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Gjodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A.Authors: Uetebach, T.; Cotton, M.D.; Spriggs, T.; Artlisch, P.; Kaime, B.P.; Sykes, S.

Smith, H.O.; Woske, C.R.; Venter, J.C.

A.Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A.Reference number: A69250; MUID:98049343; PMID:9389475

A.Accession: B69461

A.Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-420 <KLE>

A.Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PIDD:AAB8954.1; PID:9264885

C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 17.5%; Score 370.5; DB 2; Length 420;  
 Best Local Similarity 27.3%; Pred. No. 2.3e-22;  
 Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 15 VDPWTQTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDN 74  
 DB 3 VTPWEVEGV---IDYSKLEERGMQRP-SEVLPEID---NPHILMRGALFGHRDY 52  
 QY 75 QVLDAVENKKPPLYLTGSGPSSSEAMHVGHLIFFTKMLQDVFNPLVYIQMDDEKYLK 134  
 DB 53 RILEAKQKEPMAVMSGFMPSG-LPHFGHMTMDELIVMQSAGKAFV-AIDMEHNSVR 110  
 QY 135 DLTLDQAVGADAVENADIIACGFDINKTIFSDLDYMGSSGFFKAVV-KIQGHVTFNOY 193  
 DB 111 GSWETRELMGLYISIIALGREDVAIVYFOS-----KSHVHKLAFELSAVNPSEL 164  
 QY 194 KGIFGFTDSDCTGKISFPALIQAPSPNSFPQIFRDRDIOCLIPCAIDODPFRRTRDV 253  
 DB 165 RALYGNSTSLAKMVTYIAQADIL--HPQLSDGGRKPVVYVGADODPHMRILRDL 221  
 QY 254 APRI----- 257  
 DB 222 AARISIFSEPEVGVVRSRKGAELISLRLDPDKIYEEHMDIFGEABETIERAVRKI 281  
 QY 258 -----GYPKPALHSTFPFALQAGQOTKMSADPNSSIFLDTAOKITKYNKAFSGG 310  
 DB 282 EWEIGGFAPIIPSTYHRTTGLTG- -KWSSSKPESYISLDPPEGAKKVK- AFTGG 338  
 QY 311 RDTIEHRQFGNCNDVVSFMYLTFLEDD-DKLEQIRKDYTSGAMLTGELKKALIEVLQ 369

DB 339 RATAEGRRLGCEPRRCVFEIYSHFLIDSDBELNQIEHECEGNLLCKCKKMAALVK 398  
 QY 370 FLIAEHQARRKEV 382  
 DB 399 SFLKEHQRKMEAV 411

#### RESULT 15

D72477

probable tryptophanyl-tRNA synthetase ABE2461 - Aeropyrum pernix (strain K1)

C.Species: Aeropyrum pernix

C.Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C.Accession: D72477

R.Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Kaikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ka

DNA Res. 6, 83-101, 1999

A.Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A.Reference number: A72450; MUID:99310339; PMID:10382966

A.Accession: D72477

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-374 <KAW>

A.Cross-references: DDBJ:AP000064; NID:95105945; PIDD:BAA81476.1; PID:95106165

A.Experimental source: strain K1

C.Genetics:

C.Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.7%; Score 353; DB 2; Length 374;  
 Best Local Similarity 30.4%; Pred. No. 5.2e-21;  
 Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 15 VDPWTQTSAAKIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFFSHRDN 74  
 DB 8 LDPW-----GAVIKQYDRLRTFGIRPSEVL--PLRKAGMPSFLMRGIIFGHRDP 61  
 QY 75 QVLDAVENKKPPLYLTGSGPSSSEAMHVGHLIFFTKMLQ-DVFNPLVYIQMDDEKYLK 133  
 DB 62 KLEKAKGERVAVLTFGPSK-PHFGKLTVDOLIVQKNGFV--FVALDAEAFV 118  
 QY 134 KDLTDQAVGADAVEN-AKDIACGFDINKT-FISDLDYMGSSGFFKAVV-KIQGHVTFN 191  
 DB 119 RIGGEENRIVAEYIANMILGIDPKDTEYFG-----TNRGTYYFLIQFSGKVT 174  
 QY 192 QVKGIFG-FTDSDCTGKISFPALIQAPSPNSFPQIFRDRDIOCLIPCAIDODPFRRMT 250  
 DB 175 EWEATYIGELTPAKMMAISLT---QAADILHVQLDVEGYR--HVVVPVGADODPHLRILT 227  
 QY 251 RVNAPR-----IGYPRPALHSTFPFALQAGQOTKMSADPNSSIFLDTAOKITKYNK 306  
 DB 228 RDLADRMAGVLELREPASTYHNLQGLDG--RKMSSSRDSTIFLDPPEVARNLFR-A 284  
 QY 307 PSGRDITIEHRQFGNCNDV-DVSFMYLTFLEDDKLEQIRKDYTS---GAMLTGELKK 362  
 DB 285 LTGGRATAEGRRLGQVGEVCSYVHMDLHMLPDDGEVYKHT---YTSRLGKILGCECKQ 341  
 QY 363 ALIEVLQPLIAEHQARRKEVTDIVKEFTMPR 394  
 DB 342 IAMEKLERFLAEHQRLERAKTIAMKLVEPPR 373

Search completed: July 10, 2003, 12:32:18  
 Job time : 16.5539 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:30:15 ; Search time 20.4145 Seconds

(without alignments)  
2286.817 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGDPATEAEEDFVDPMTV.....VTDEIVKEFMPKISFPDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	415	9	US-09-813-718-14
2	2116	100.0	437	9	US-09-813-718-12
3	2116	100.0	471	9	US-10-126-467B-2
4	2116	100.0	484	9	US-09-813-718-10
5	2101	99.3	471	9	US-09-919-039-163
6	2101	99.3	475	10	US-09-925-302-558
7	1988	94.0	392	9	US-09-813-718-16
8	1125	53.2	433	9	US-10-128-714-8545
9	831	39.3	173	10	US-09-925-302-855
10	414.5	19.6	179	9	US-10-128-714-3545
11	328	15.5	85	9	US-09-813-718-45
12	292	13.8	85	9	US-09-813-718-46
13	273.5	12.9	85	9	US-09-813-718-48
14	263	12.4	85	9	US-09-813-718-47
15	226.5	10.7	142	10	US-09-925-302-557
16	185.5	8.8	341	10	US-09-815-242-13444
17	147	6.9	46	9	US-09-813-718-51
18	147	6.9	385	9	US-10-128-714-3379
19	146.5	6.9	391	9	US-10-128-714-8379

20	146	6.9	339	10	US-09-815-242-11422	Sequence 11422, A
21	142.5	6.7	337	9	US-10-156-761-10954	Sequence 10954, A
22	134	6.3	372	9	US-09-813-718-8	Sequence 8, Appl1
23	132.5	6.3	372	9	US-09-813-718-4	Sequence 4, Appl1
24	132.5	6.3	372	9	US-09-813-718-2	Sequence 2, Appl1
25	128	6.0	334	10	US-09-815-242-11070	Sequence 11070, A
26	120	5.7	46	9	US-09-813-718-52	Sequence 52, Appl1
27	113.5	5.4	423	10	US-09-815-242-10678	Sequence 10678, A
28	113	5.3	415	10	US-09-815-242-5494	Sequence 5494, Ap
29	113	5.3	420	10	US-09-815-242-12649	Sequence 12649, A
30	112	5.3	345	9	US-09-738-626-4252	Sequence 4252, Ap
31	110.5	5.2	331	9	US-10-156-761-12259	Sequence 12259, A
32	108	5.1	306	9	US-10-126-927-50	Sequence 50, Appl1
33	108	5.1	306	9	US-10-126-927-59	Sequence 59, Appl1
34	108	5.1	306	9	US-10-126-927-64	Sequence 64, Appl1
35	108	5.1	306	9	US-10-126-931A-50	Sequence 50, Appl1
36	108	5.1	306	9	US-10-126-931A-59	Sequence 59, Appl1
37	108	5.1	306	9	US-10-126-931A-64	Sequence 64, Appl1
38	108	5.1	334	10	US-09-815-242-13806	Sequence 13806, A
39	107	5.1	306	9	US-10-126-927-51	Sequence 51, Appl1
40	107	5.1	306	9	US-10-126-931A-51	Sequence 51, Appl1
41	106.5	5.0	334	10	US-09-815-242-10371	Sequence 10371, A
42	106	5.0	306	9	US-10-126-927-45	Sequence 45, Appl1
43	106	5.0	306	9	US-10-126-931A-45	Sequence 45, Appl1
44	105.5	5.0	46	9	US-09-813-718-54	Sequence 54, Appl1
45	105	5.0	306	9	US-10-126-927-43	Sequence 43, Appl1

## ALIGNMENTS

RESULT 1  
US-09-813-718-14  
Sequence 14, Application US/09813718  
Publication No. US2002018266A1  
GENERAL INFORMATION:  
APPLICANT: Schimmel, Paul  
APPLICANT: Makasugi, Keiuke  
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
TITLE OF INVENTION: The Regulation of Angiogenesis  
FILE REFERENCE: 00-221  
CURRENT APPLICATION NUMBER: US/09/813,718  
CURRENT FILING DATE: 2001-03-21  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 415  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human  
OTHER INFORMATION: supermatl tips in pET20B  
US-09-813-718-14

Query Match	100.0%	Score 2116;	DB 9;	Length 415;
Best Local Similarity	100.0%	Pred. No. 5.7e-192;		
Matches 401;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	SNHGDPATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIEATQGRPH 60			
2	SNHGDPATEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIEATQGRPH 61			
61	FLRRGIFPSHRMNOVLAYENKKPFYLYTGSPSEAMHVGHLPIFTKMLQVFNVP 120			
62	FLRRGIFPSHRMNOVLAYENKKPFYLYTGSPSEAMHVGHLPIFTKMLQVFNVP 121			
121	LVIQMTDEKYLKLTLDQAYGAVENAKDIIACGFDINKTIFPSDIDYMGSSGFYKN 180			
122	LVIQMTDEKYLKLTLDQAYGAVENAKDIIACGFDINKTIFPSDIDYMGSSGFYKN 181			
181	VVKIQHVTENVKGI FGFSDICIGKISFPALQAAPSPNSFPQIFRDRTDIOCLIPCA 240			
182	VVKIQHVTENVKGI FGFSDICIGKISFPALQAAPSPNSFPQIFRDRTDIOCLIPCA 241			

```

QY 241 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 300
DB 242 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 301
QY 301 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 302 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 361
QY 361 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 401
DB 362 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 402

```

```

RESULT 2
US-09-813-718-12
; Sequence 12, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 437
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human mini
; OTHER INFORMATION: Trps in per208
US-09-813-718-12

```

```

Query Match 100.0%; Score 2116; DB 9; Length 437;
Best Local Similarity 100.0%; Pred. No. 6.1e-192;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SNHGPDATAEEDFVDPMVTQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 60
DB 24 SNHGPDATAEEDFVDPMVTQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 83
QY 61 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 120
DB 84 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 143
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 144 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203
QY 181 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALQGAQTSASPNSSIFLTTAKQIKT 240
DB 204 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALQGAQTSASPNSSIFLTTAKQIKT 263
QY 241 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 300
DB 264 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 323
QY 301 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 324 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 383
QY 361 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 401
DB 384 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 424

```

```

RESULT 3
US-10-126-467B-2
; Sequence 2, Application US/10126467B

```

```

; Publication No. US20030059797A1
; GENERAL INFORMATION:
; APPLICANT: Paley, Elena
; TITLE OF INVENTION: ANIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE
; FILE REFERENCE: PAL-111
; CURRENT APPLICATION NUMBER: US/10/126,467B
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/284,980
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 09/513,895
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 09/384,869
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 471
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-126-467B-2

```

```

Query Match 100.0%; Score 2116; DB 9; Length 471;
Best Local Similarity 100.0%; Pred. No. 6.8e-192;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SNHGPDATAEEDFVDPMVTQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 60
DB 71 SNHGPDATAEEDFVDPMVTQTSASAKIDYDKLIVRFGSSKIDKELINRIERATGQRPH 130
QY 61 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 120
DB 131 FLRRGIFFSHRDNNQVLDAVENKKPFYLYTGRGSSAMHVGHLIPFTKMLQDVFNVP 190
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
QY 181 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALQGAQTSASPNSSIFLTTAKQIKT 240
DB 251 VVKIQKAVTFNQVKGIFGFTDSDCIKISFPALQGAQTSASPNSSIFLTTAKQIKT 310
QY 241 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 300
DB 311 IDDDPYFRMTRDVAAPRIGYKPPALHSTFFPALQGAQTQMSADPNSSIFLTTAKQIKT 370
QY 301 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 371 KVNKAHAFSGGRDTEEHROFGNCDDVVSFWYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 401
DB 431 KKALIEVLQPLIAEHQARRKEVTDIVKEFMTPRKLSFDFQ 471

```

```

RESULT 4
US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 464
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Description of Artificial Sequence: human  
OTHER INFORMATION: full-length Trps in pET20B  
US-09-813-718-10

Query Match 100.0%; Score 2116; DB 9; Length 484;  
Best Local Similarity 100.0%; Pred. No. 7.1e-192;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60  
DB 71 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 130  
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120  
DB 131 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190  
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 180  
DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 250  
QY 181 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 240  
DB 251 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 310  
QY 241 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 300  
DB 311 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 370  
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 360  
DB 371 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 430  
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401  
DB 431 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 471

## RESULT 5

US-09-919-039-163  
Sequence 163, Application US/09919039  
Publication No. US20030108871A1  
GENERAL INFORMATION:  
APPLICANT: Kaser, Matthew R.  
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
FILE REFERENCE: PA-0035 US  
CURRENT APPLICATION NUMBER: US/09/919, 039  
CURRENT FILING DATE: 2002-09-09  
PRIOR APPLICATION NUMBER: 60/222, 113  
PRIOR FILING DATE: 2000-07-28  
NUMBER OF SEQ ID NOS: 401  
SOFTWARE: PERL Program  
SEQ ID NO 163  
LENGTH: 471  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030108871A1 2705515CD1  
US-09-919-039-163

Query Match 99.3%; Score 2101; DB 9; Length 471;  
Best Local Similarity 99.5%; Pred. No. 1.8e-190;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60  
DB 71 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 130  
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120  
DB 131 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190  
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 180

DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 250  
QY 181 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 240  
DB 251 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 310  
QY 241 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 300  
DB 311 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 370  
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 360  
DB 371 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 430  
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401  
DB 431 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 471

## RESULT 6

US-09-925-302-558  
Sequence 558, Application US/09925302  
Patent No. US20020044941A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA104  
CURRENT APPLICATION NUMBER: US/09/925, 302  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05918  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124, 270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 896  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 558  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-302-558

Query Match 99.3%; Score 2101; DB 10; Length 475;  
Best Local Similarity 99.5%; Pred. No. 1.8e-190;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 60  
DB 75 SNHGDATEAEEDFVDPMTVOTSSAKGIDYDKLIVRFGSSKIDKELINRERATGQRPH 134  
QY 61 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120  
DB 135 FLRRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 194  
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 180  
DB 195 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINTKPIFSDLDYMGSSGFYKN 254  
QY 181 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 240  
DB 255 VKIKQHTVFNQVKIGFETSDSDCIKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 314  
QY 241 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 300  
DB 315 IDDDYFFMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSADPNSSIFLTDTKAQIKT 374  
QY 301 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 360  
DB 375 KVNKAIFSGGRDTIEHRQFGNCDVDVSPMYLTFELEDDBDLBOIRKDYTSGAMLTGEL 434  
QY 361 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401  
DB 435 KKALEIVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 475

```
RESULT 7
US-09-813-718-16
; Sequence 16, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakauegi, Keisuke
; TITLE OF INVENTION: Human Aminocycl-1-erNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813, 718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.0
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: human minor
; OTHER INFORMATION: Trps fragment in pET20B
US-09-813-718-16

Query Match          94.0%; Score 1988; DB 9; Length 392;
Best Local Similarity 100.0%; Pred. No. 6,9e-180;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 SAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHHFRGIFFSHRDNNQVLDAYENK 83
DB 2 SAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHHFRGIFFSHRDNNQVLDAYENK 61
QY 84 KPEYLYTGRSPSSSEAMHVGHLIPFTFKWLQDVFNVLVQMTDDEKYLTKDLTDQAYG 143
DB 62 KPEYLYTGRSPSSSEAMHVGHLIPFTFKWLQDVFNVLVQMTDDEKYLTKDLTDQAYG 121
QY 144 DAVENAKDIIACGPDINKTFIFSDLDYMGSSGKYKVVVKIOKHVFNQVKGIFGFSDSD 203
DB 122 DAVENAKDIIACGPDINKTFIFSDLDYMGSSGKYKVVVKIOKHVFNQVKGIFGFSDSD 181
QY 204 CIKGISFPALQAPSPFSNSFPQIFRDRDIDQCLIPCAIDDPYFRMTRDVAAPRIGYKPA 263
DB 182 CIKGISFPALQAPSPFSNSFPQIFRDRDIDQCLIPCAIDDPYFRMTRDVAAPRIGYKPA 241
QY 264 LHMSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVKKHAFSGGRDITIEHRQFGCN 301
DB 242 LHMSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVKKHAFSGGRDITIEHRQFGCN 301
QY 324 CDVDVSPFMYLTFFLEDDDKLEQIRKDYSGAMLTGELKKALIEVLOPLIAEHQARRKEVT 383
DB 302 CDVDVSPFMYLTFFLEDDDKLEQIRKDYSGAMLTGELKKALIEVLOPLIAEHQARRKEVT 361
QY 384 DEIVKEFMTPRKLSFDFQ 401
DB 362 DEIVKEFMTPRKLSFDFQ 379

RESULT 8
US-10-128-714-8545
; Sequence 8545, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroszhin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
```

```
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 8545
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8545

Query Match          53.2%; Score 1125; DB 9; Length 433;
Best Local Similarity 54.7%; Pred. No. 4,2e-98;
Matches 220; Conservative 63; Mismatches 107; Indels 12; Gaps 5;

QY 7 ATEAEEDFYDPMYTV-----QTSASAKGIDYDKLIVRFGSSKIDKELINRTERATGCRPHH 61
DB 22 ASKAVAQVYTPFDVSGVDSGKLPVDYDKLVREGRGARISKELEFEREYTGRRPHH 81
QY 62 LRRGIFFSHRDNNQVLDAYENKPPYLYTGRSPSSSEAMHVGHLIPFTFKWLQDVFNVL 121
DB 82 MRRGIFFSHRDNLILIDRYEKQPFYLYTGRSPSSSEAMHVGHLIPFTFKWLQDVFNVL 141
QY 122 VIQMTDDEKYLTKDLTDQAYGDAVENAKDIIACGPDINKTFIFSDLDYMGSSGKYK 180
DB 142 VIQMTDDEKYLTKDLTDQAYGDAVENAKDIIACGPDINKTFIFSDLPVGG--GAFYEN 199
QY 181 VVKIQKHVFNQVKGIFGFSDSDCIKGISFPALQAPSPFSNSFPQIF--RDR--TDIOCL 236
DB 200 ICMARILINISRGIFGNDNNVEHFCAQTQSTATATSPHIFGDRKKVSSIPCL 259
QY 237 IPCAIDDPYFRMTRDVAAPRIGYKPAALHSTFPALQAGQTKMSASDPNSSIFLDTAK 296
DB 260 IPCAIDDPYFRMTRDVAAPRIGYKPAALHSTFPALQAGQTKMSASDPNSSIFLDTAK 319
QY 297 QIKTKVKKHAFSGGRDITIEHRQFGCNCDVDVSPFMYLTFFLEDDDKLEQIRKDYSGAM 356
DB 320 RINKKTKVAFSGGDTALQROLGANTYKODVPFOYLTFFMEDDELEIRIVAYEKEM 379
QY 357 TGELEKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSF 398
DB 380 TGEVKKCICIAELQAYVQAFQERRAQVTDEIVAEFMRPSLEW 421

RESULT 9
US-09-925-302-855
; Sequence 855, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 855
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: SITE  
LOCATION: (159)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (168)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-302-855

Query Match 39.3%; Score 831; DB 10; Length 173;  
Best Local Similarity 98.7%; Pred. No. 7.9e-71;

Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 107 F1FTKMLQOVENVPLVYQMTDEKYLKDLTDQAVGADVENAKDIACGFDINKTFIRS 166  
DB 1 F1FTKMLQOVENVPLVYQMTDEKYLKDLTDQAVGADVENAKDIACGFDINKTFIRS 60  
QY 167 DLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIGKISFPAIQAAFSFNSFPQI 226  
DB 61 DLDYMGSSGFYKNNVKIQKHTFNQVKGIFGFTSDCIGKISFPAIQAAFSFNSFPQI 120  
QY 227 FRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKAL 264  
DB 121 FRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKAL 158

RESULT 10  
US-10-128-714-3545

Sequence 3545, Application US/10128714  
Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo

APPLICANT: Hu, Weng

APPLICANT: Tishkoff, Daniel

APPLICANT: Zamudio, Carlos

APPLICANT: Broshkin, Alexey M

APPLICANT: Lemieux, Sebastien M

TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

FILE REFERENCE: 10182-018-999

CURRENT APPLICATION NUMBER: US/10/128, 714

CURRENT FILING DATE: 2002-04-23

PRIOR APPLICATION NUMBER: US 60/285,697

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: US 60/287,066

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/295,890

PRIOR FILING DATE: 2001-06-05

PRIOR APPLICATION NUMBER: US 60/303,899

PRIOR FILING DATE: 2001-07-09

PRIOR APPLICATION NUMBER: US 60/316,362

PRIOR FILING DATE: 2001-08-31

NUMBER OF SEQ ID NOS: 8603

SOFTWARE: Patent in version 3.1

SEQ ID NO 3545

LENGTH: 179

TYPE: PR

ORGANISM: Aspergillus fumigatus

US-10-128-714-3545

Query Match 19.6%; Score 414.5; DB 9; Length 179;  
Best Local Similarity 59.4%; Pred. No. 2.3e-11;

Matches 76; Conservative 20; Mismatches 27; Indels 5; Gaps 1;

QY 7 ATAEAEEDFVDPMTV-----QTSAGKIDYDKLIVRFSSSKIDKELINRIERATGCRPHH 61  
DB 22 ASRAVAQVVTTPFDVSGVDSGLLPVDYDKLIVREAGATRISEKELLEPEREVTGRRPHR 81  
QY 62 LRGGIFFSHRDMNQVDAVENKKPFYLYGGRGSSSEAMVGHILPFTKMLQOVENVPL 121  
DB 82 MRGIGVFSHRDLNLIIDREKGGPFYLYGGRGSSSEAMVGHITPFTKMLQOVENVPL 141  
QY 122 VIQMTDE 129  
|||

DB 142 VIQMTDE 149

RESULT 11

US-09-813-718-45

Sequence 45, Application US/09813718  
Publication No. US2002018266A1

GENERAL INFORMATION:

APPLICANT: Schimmel, Paul

APPLICANT: Wakasugi, Keisuke

TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For

FILE REFERENCE: 00-221

CURRENT APPLICATION NUMBER: US/09/813,718

CURRENT FILING DATE: 2001-03-21

NUMBER OF SEQ ID NOS: 58

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 45

LENGTH: 85

TYPE: PR

ORGANISM: Homo sapiens

US-09-813-718-45

Query Match 15.5%; Score 328; DB 9; Length 85;  
Best Local Similarity 100.0%; Pred. No. 1.2e-23;

Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPMTVQTSAGKIDYDKLIVRFSSSKIDKELINRIERATGCRPHH 60  
DB 24 SNHGPDATAEEDFVDPMTVQTSAGKIDYDKLIVRFSSSKIDKELINRIERATGCRPHH 83  
QY 61 FL 62  
DB 84 FL 85

RESULT 12

US-09-813-718-46

Sequence 46, Application US/09813718  
Publication No. US2002018266A1

GENERAL INFORMATION:

APPLICANT: Schimmel, Paul

APPLICANT: Wakasugi, Keisuke

TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For

FILE REFERENCE: 00-221

CURRENT APPLICATION NUMBER: US/09/813,718

CURRENT FILING DATE: 2001-03-21

NUMBER OF SEQ ID NOS: 58

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 46

LENGTH: 85

TYPE: PR

ORGANISM: Bos taurus

US-09-813-718-46

Query Match 13.8%; Score 292; DB 9; Length 85;  
Best Local Similarity 90.3%; Pred. No. 3.2e-20;

Matches 56; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPMTVQTSAGKIDYDKLIVRFSSSKIDKELINRIERATGCRPHH 60  
DB 24 SGEGLDATABEDFVDPMTVQTSAGKIDYDKLIVRFSSSKIDKELINRIERATGCRPHH 83  
QY 61 FL 62  
DB 84 FL 85  
RESULT 13  
US-09-813-718-48  
Sequence 48, Application US/09813718  
Publication No. US2002018266A1

```
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; TITLE OF INVENTION: Human Antinocyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 48
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-813-718-48

Query Match      12.9%; Score 273.5; DB 9; Length 85;
Best Local Similarity 83.9%; Pred. No. 1.8e-18;
Matches 52; Conservative 6; Mismatches 3; Indels 1; Gaps 1;

QY      2 NHGPD-ATEABEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60
Db      24 SHGDEAVDDEDFVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 83
QY      61 FL 62
Db      84 FL 85

RESULT 14
US-09-813-718-47
; Sequence 47, Application US/09813718
; Publication No. US2002018266A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Antinocyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-813-718-47

Query Match      12.4%; Score 263; DB 9; Length 85;
Best Local Similarity 85.2%; Pred. No. 1.8e-17;
Matches 52; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 NHGPD-ATEABEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 61
Db      25 NCSDATKASEDFVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 84
QY      62 L 62
Db      85 L 85

RESULT 15
US-09-925-302-557
; Sequence 557, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
```

```
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 557
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (130)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-557

Query Match      10.7%; Score 226.5; DB 10; Length 142;
Best Local Similarity 70.6%; Pred. No. 1.1e-13;
Matches 48; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY      1 SNHGP-ATEABEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 60
Db      75 SNHGP-ATEABEDPVDPTVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRP 133
QY      61 FLRGIFF 68
Db      134 --GQGXFF 139
```

Search completed: July 10, 2003, 12:33:37  
Job time: 21.4145 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:29:05 ; Search time 12.1515 Seconds  
(without alignments)  
970.956 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471  
Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPWTV.....VTDEIVKERTPKLSFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents AA:\*  
2: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	57.6	424	4	US-08-876-885-26 Sequence 26, Appl
2	185.5	8.8	341	3	US-08-928-100-2 Sequence 2, Appl
3	185.5	8.8	341	4	US-09-492-581-2 Sequence 2, Appl
4	185.5	8.8	341	4	US-09-425-666-2 Sequence 2, Appl
5	164.5	7.8	409	2	US-08-743-130A-39 Sequence 39, Appl
6	162.5	7.7	409	2	US-08-743-130A-2 Sequence 2, Appl
7	132.5	6.3	388	3	US-08-705-868-4 Sequence 4, Appl
8	132.5	6.3	388	3	US-09-123-615-4 Sequence 4, Appl
9	113.5	5.4	418	4	US-08-855-910-11 Sequence 11, Appl
10	109	5.2	377	4	US-09-352-990-28 Sequence 28, Appl
11	108	5.1	197	3	US-08-923-867-2 Sequence 2, Appl
12	108	5.1	197	3	US-08-928-100-4 Sequence 4, Appl
13	108	5.1	197	4	US-09-183-134-2 Sequence 4, Appl
14	108	5.1	197	4	US-09-492-581-4 Sequence 4, Appl
15	108	5.1	197	4	US-09-425-666-4 Sequence 4, Appl
16	104.5	4.9	370	2	US-08-415-593-45 Sequence 45, Appl
17	100.5	4.7	427	4	US-09-134-001C-5141 Sequence 5141, Ap
18	95	4.5	418	3	US-08-844-054-2 Sequence 2, Appl
19	95	4.5	418	4	US-09-347-333-2 Sequence 2, Appl
20	95	4.5	464	4	US-09-134-001C-4701 Sequence 4701, Ap
21	94.5	4.5	877	2	US-08-907-166-8 Sequence 8, Appl
22	89.5	4.2	344	4	US-09-393-554-2 Sequence 2, Appl
23	88.5	4.2	501	4	US-09-157-257-8 Sequence 8, Appl
24	88.5	4.2	898	1	US-08-465-995A-4 Sequence 4, Appl
25	88.5	4.2	898	2	US-08-465-994C-4 Sequence 4, Appl
26	88.5	4.2	898	2	US-08-966-145-4 Sequence 4, Appl
27	88.5	4.2	920	1	US-08-101-593-4 Sequence 4, Appl

28	88	4.2	606	2	US-08-883-534-3 Sequence 3, Appl
29	88	4.2	606	3	US-09-204-764-3 Sequence 3, Appl
30	87.5	4.1	855	4	US-08-890-865A-10 Sequence 10, Appl
31	87.5	4.1	898	1	US-08-465-995A-2 Sequence 2, Appl
32	87.5	4.1	898	2	US-08-465-994C-2 Sequence 2, Appl
33	87.5	4.1	898	2	US-08-966-145-2 Sequence 2, Appl
34	87.5	4.1	920	1	US-08-101-593-2 Sequence 2, Appl
35	87	4.1	502	4	US-09-134-001C-4511 Sequence 4511, Ap
36	86.5	4.1	849	4	US-09-157-257-4 Sequence 4, Appl
37	86	4.1	930	4	US-09-134-001C-5314 Sequence 5314, Ap
38	85	4.0	410	4	US-09-352-990-16 Sequence 16, Appl
39	84.5	4.0	539	3	US-09-157-257-6 Sequence 6, Appl
40	83.5	3.9	428	3	US-08-331-625A-43 Sequence 43, Appl
41	83.5	3.9	428	4	US-09-494-153-43 Sequence 43, Appl
42	83.5	3.9	510	1	US-08-249-112-3 Sequence 3, Appl
43	83.5	3.9	510	5	PCT-US95-06556-3 Sequence 3, Appl
44	83.5	3.9	970	1	US-08-375-709-7 Sequence 7, Appl
45	83.5	3.9	970	1	US-08-752-929-7 Sequence 7, Appl

#### ALIGNMENTS

RESULT 1

US-08-876-885-26  
Sequence 26, Application US/08876885  
Patent No. 6174713

GENERAL INFORMATION:

APPLICANT: Shen, Xiaoyu

APPLICANT: Houman, Fariba

TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-tRNA

TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING

TITLE OF INVENTION: SAME

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/876,885

FILING DATE: 16-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: CPI97-02

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-876-885-26

Query Match 57.6%; Score 1218.5; DB 4; Length 424;  
Best Local Similarity 57.1%; Pred. No. 3.7e-128;  
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;  
QY 8 TEAEDFVDPWTVQ---TSSAKGIDYKLVFGSSKIDKELNIRIRATGQPHHPLR 63  
DB 13 TEESQKITPWEVEGAVVDGKMGIDYDKLISQFGTKHITETTLERFQVGTGEEDPFLAK 72

QY 64 RGTFESHDMNVLDAYENKKPFYLYTGRGSSSEAMVGHLPFTKMLQDVENVPLVI 123  
DB 73 RGFFSRDLDRIIDLVEHGFPLTYGRGSSSMHGHVPIFTKMLQDFVPLVI 132  
QY 124 QMTDEKYLWK-DLTLDQAYGDAVENAKDIJACGFDINKTFIFSDLDYMGSSGFYKNV 182  
DB 133 ELTDEKFLFPHQTLTIDVKGFAENAKDIJAVGFNPENFTFISDLQYMG--GAFYENVV 190  
QY 183 KICKHVFNOVKIGFETDSDCTGKISFPALQAAFPSNSFPQIFRRTDIQCLIPCAID 242  
DB 191 RTRQITTSYAKAVGFTDSDCTGKIHFASTQIATAFPSSPDVGLPPTKPLIPCAID 250  
QY 243 QDVFYRTRDVAPRIGYKPKALLHSTFPALQAGQTKMSADPNSSIFLDTAKQIKTV 302  
DB 251 QDVFYRTRDVADKLTFTKALITAKFPALQAGSTMSADTTSTIFMGDTAKQIKTI 310  
QY 303 NKHAFSGGRDTIEHRQFGNCVDVSPMYLTFLEDDDKLEQIRKDYTSGAMLTGELKK 362  
DB 311 NKYAFSGGRATAEHRRELGNPEVDVAFQYLSFYSYDEKLAQLQEGYRKGEILSGMKK 370  
QY 363 ALIEVLOPLAEHQARKVETDEIVKEFMTPKLSF 398  
DB 371 ECITVLOEFVSAQERSKVDQVEKEMKPKLVF 406

## RESULT 2

US-08-928-100-2  
Sequence 2, Application US/08928100

Patent No. 6046174  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6046174el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-2

Query Match 8.8%; Score 185.5; DB 3; Length 341;  
Best Local Similarity 24.0%; Pred. No. 3,2e-12;

Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;  
QY 84 KPPYLYTGKGPSSSEAMVGHLPFTKMLQDVENVPLVIQMTDEKY-LMKDULTLDQAY 142  
DB 3 KPYIL-TGDRPTOK-HIGHYVGSILNR-----VLQDEBKYMFPFLADQOAL 49  
QY 143 GDIVEN-----AKDIJACGFDINKTFIF--SDLDYMGSSGFYKNV---KICK 186  
DB 50 TDHAKPQITVBSIGNVALDYLAAGDPKSTIFITOSQIPELAEISMYNVLSTLRLR 109  
QY 187 HTFVNOVKIGFETDSDCTGKISFPALQAA--PSFNSFPQIFRRTDIQCLIPCAID 244  
DB 110 NPTVKTEISQKGEISIPGFLVYPIAQADITAFKANY-----VPVGTQDK 156  
QY 245 PYRMTRD-----VAPRIGYKPKALLHSTFPALQAGQTKMSADPNSSIF 290  
DB 157 PMIQREIRIVRSFNANVCDVLEPPIYFENE--RAGRLPGIDG--NAKMSKS-LNNGIY 212  
QY 291 LDTDAKQIKTVKNKHAFAFGGRDTIEHRQFGNCVDVSPMYLTFE--LEDDDKLEQIRK 348  
DB 213 LADDAITLKKVMSMTTDDHIREVDPKIEGN---WVFHLDVGRFEDPAQELADMK 268  
QY 349 DYTSGAMLTGELKALIEVLOPLAEHQARKVETDEI 386  
DB 269 RYORGGLGVKTKRYLLEIRERLGPFRRIEFKDM 306

## RESULT 3

US-09-492-581-2  
Sequence 2, Application US/09492581

Patent No. 6346409  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6346409el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/492,581  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
APPLICATION NUMBER: 9619072.3  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein



US-09-492-581-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
Best Local Similarity 24.0%; Pred. No. 3.2e-12;  
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPEVLTGSGPSSSEAMHVGHLIPFTKMLQDVFNPLVLTQMTDDEKY-LMKDLTLDDQAY 142  
DB 3 KPILL-TGDRPTCK-LHIGHVGSLSKNR-----VLLQEBKQDMFVFLADQOAL 49  
QY 143 GDAVEN-----AKDIIACGFDINKTFIF--SDLDYMGSSGFYKXNV---KIQK 186  
DB 50 THAKDPQITVIESIGNVALDYLAAGLDPNKSTIFIQSQIPELAEISMTYMLVSLARLER 109  
QY 187 HTFENQVKGIFFGPTSDCIGKISFPALQAA--PSFNSFPQIFRDRTDIOCLIPCAIDOD 244  
DB 110 NPTVKTETISOKGFGSIPGFLVYPLAQADITAFKANY-----VPVGTQDK 156  
QY 245 PYFRMTRD-----VAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIF 290  
DB 157 PMIEQIREIVRSFNNAVNCVIVPEPGIYPENE--RAGRLPGIDG-NAMKSKS-LNGIY 212  
QY 291 LDDTAQKQITKYNKAHAFSGGRDTIEEHROFGNCVDVSEFWYLTFF--LEDDKLEQIRK 348  
DB 213 LADDADTLRKKWSMTTDPDHIRVEDPGKIEGN---WVFHYLDVFGREDPAQEIADMKR 268  
QY 349 DYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDEI 386  
DB 269 RYRGGLGDVTKRYLLEILERELGPIRRRIEFPAKM 306

## RESULT 4

US-09-425-666-2

; Sequence 2, Application US/09425666  
; Patent No. 6416976

GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6416976el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/425,666  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/928,100  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 341 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-425-666-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
Best Local Similarity 24.0%; Pred. No. 3.2e-12;  
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPEVLTGSGPSSSEAMHVGHLIPFTKMLQDVFNPLVLTQMTDDEKY-LMKDLTLDDQAY 142  
DB 3 KPILL-TGDRPTCK-LHIGHVGSLSKNR-----VLLQEBKQDMFVFLADQOAL 49  
QY 143 GDAVEN-----AKDIIACGFDINKTFIF--SDLDYMGSSGFYKXNV---KIQK 186  
DB 50 THAKDPQITVIESIGNVALDYLAAGLDPNKSTIFIQSQIPELAEISMTYMLVSLARLER 109  
QY 187 HTFENQVKGIFFGPTSDCIGKISFPALQAA--PSFNSFPQIFRDRTDIOCLIPCAIDOD 244  
DB 110 NPTVKTETISOKGFGSIPGFLVYPLAQADITAFKANY-----VPVGTQDK 156  
QY 245 PYFRMTRD-----VAPRIGYKPKALHSTFPALQAGQTKMSASDPNSSIF 290  
DB 157 PMIEQIREIVRSFNNAVNCVIVPEPGIYPENE--RAGRLPGIDG-NAMKSKS-LNGIY 212  
QY 291 LDDTAQKQITKYNKAHAFSGGRDTIEEHROFGNCVDVSEFWYLTFF--LEDDKLEQIRK 348  
DB 213 LADDADTLRKKWSMTTDPDHIRVEDPGKIEGN---WVFHYLDVFGREDPAQEIADMKR 268  
QY 349 DYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDEI 386  
DB 269 RYRGGLGDVTKRYLLEILERELGPIRRRIEFPAKM 306

## RESULT 5

US-08-743-130A-39

; Sequence 39, Application US/08743130A  
; Patent No. 5871987

GENERAL INFORMATION:  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Tao, Nianjun  
APPLICANT: Tao, Jianshi  
TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,130A  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPI95-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids

TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-743-130A-39

Query Match 7.8%; Score 164.5; DB 2; Length 409;  
Best Local Similarity 20.7%; Pred. No. 1e-09;  
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAYENK-KPFLYLTGRGSSSEAMVHGLIPFI-----FTKWLQDV-----F 117  
DB 27 IKDVLKERNRPVXIYWGTAFTGKP-HCGYFVPMIKLAHFLKAGCEVTLADLHAFLDNM 85  
QY 118 NVPLVQMTDDEKYLKMDLTLDOAYGDAVENAKDIIACGFDINKTIFISDLDYMGSSGF 177  
DB 86 KALEVVKYKRAKYEFVVKAILKSIINVPIERLKFVVGSSYQKGDDV--MDLFKLSNIV 142  
QY 178 YKRVVK-----IQKVTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRTD 232  
DB 143 SQNDARAGADVKKVQANPLLSGLI-----YPLMQA-----IDEEHLG 180  
QY 233 IQCLIPCAIDQDPYFEMTRDVAPRIGYKPKALHSTFPALQAGQTMASDPNSIFLT 292  
DB 181 VDAQFG-GVDQKIFVLAENLPSIGYKKAHLNMPVPGI-GGGKMSASDPNSKIDII 238  
QY 293 DTKAKIKTKYKNAFSGG--RDT-----IEEHROGCGN 323  
DB 239 EEPKVVKKVNSAYCAPGELKDNGLAFIEYVIQPIAELKTGVEGAFKLDIDPEKYG- 297  
QY 324 CDVDSFMYLTFLEBDDKLEQIRKDYSGAMLTGELKALI---EVLQPLIAEHQARR 379  
DB 298 ---PLSY-----DSIEQLKADPVGDKLAPPLKSGVADKINELLAPIRAEFESS- 343  
QY 380 KEVTDEIVKEFMTPRK 395  
DB 344 -----EEFQVAQK 351

RESULT 6  
US-08-743-130A-2  
Sequence 2, Application US/08743130A  
Patent No. 5871987  
GENERAL INFORMATION:  
APPLICANT: Sagesanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Tao, Mianjun  
APPLICANT: Tao, Jianhui  
APPLICANT: Houman, Fariba  
TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,130A  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CP195-12  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-743-130A-2

Query Match 7.7%; Score 162.5; DB 2; Length 409;  
Best Local Similarity 20.7%; Pred. No. 1.7e-09;  
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAYENK-KPFLYLTGRGSSSEAMVHGLIPFI-----FTKWLQDV-----F 117  
DB 27 IKDVLKERNRPVXIYWGTAFTGKP-HCGYFVPMIKLAHFLKAGCEVTLADLHAFLDNM 85  
QY 118 NVPLVQMTDDEKYLKMDLTLDOAYGDAVENAKDIIACGFDINKTIFISDLDYMGSSGF 177  
DB 86 KALEVVKYKRAKYEFVVKAILKSIINVPIERLKFVVGSSYQKGDDV--MDLFKLSNIV 142  
QY 178 YKRVVK-----IQKVTFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRTD 232  
DB 143 SQNDARAGADVKKVQANPLLSGLI-----YPLMQA-----IDEEHLG 180  
QY 233 IQCLIPCAIDQDPYFEMTRDVAPRIGYKPKALHSTFPALQAGQTMASDPNSIFLT 292  
DB 181 VDAQFG-GVDQKIFVLAENLPSIGYKKAHLNMPVPGI-GGGKMSASDPNSKIDII 238  
QY 293 DTKAKIKTKYKNAFSGG--RDT-----IEEHROGCGN 323  
DB 239 EEPKVVKKVNSAYCAPGELKDNGLAFIEYVIQPIAELKTGVEGAFKLDIDPEKYG- 297  
QY 324 CDVDSFMYLTFLEBDDKLEQIRKDYSGAMLTGELKALI---EVLQPLIAEHQARR 379  
DB 298 ---PLSY-----DSIEQLKADPVGDKLAPPLKSGVADKINELLAPIRAEFESS- 343  
QY 380 KEVTDEIVKEFMTPRK 395  
DB 344 -----EEFQVAQK 351

RESULT 7  
US-08-705-868-4  
Sequence 4, Application US/08705868  
Patent No. 5885798  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Murty, Lynn E.  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,868  
FILING DATE: Filed Herewith  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0117 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1184699  
US-08-705-868-4

Query Match 6.3%; Score 132.5; DB 2; Length 388;  
Best Local Similarity 20.8%; Pred. No. 3.6e-06;  
Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

QY 91 GRGPS-SEAMHVGHLIPFTKMLQDVFNPLVIGMTDE--KYLKDLTLDO--AYGD 144  
DB 2 GDAPSPEEKH-----LITRNLOEVLGEKELIKERELKIYWGATTKGPHVAFV 54  
QY 145 AVENAKDIIACGFDINKTFIFSDL-----DYMGSFGFYKNVYKIOKH--YTF 190  
DB 55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELERVSYENVIKAMLESIGVPL 112  
QY 191 NOVKGIFG-----FTDSD-----CIGKISFPAIOA 216  
DB 113 EKLKFKIGTDYOLSKSEYTLDVYRLSSVTOHDSKKAAGAVVQVHEHPILISGLLYPGLQ- 171  
QY 217 PSFNSFPQIFRDRTDIOCLIPCAIDODPYFRMTDVAPRIGYPPKALLHSTFFPALOGA 276  
DB 172 --LDEEYLVK-----DAQF--GGIDQKIFTFPAEKYLPALGYSKRVHLMNPMVGLTG- 220  
QY 277 QTKMSASDPNSSIFLTDPAKQIKTKVKNKHAFA-----SGGRDTIEH----- 317  
DB 221 -SKMSSSEESKIDLLDRKEDYKKLKK-AFCBPGVNNNGVLSPIKHYLPPLKSEFVIL 278  
QY 318 --RQFGNCDVDVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLIAEH 375  
DB 279 RDEKMGKN-----KTYTAYVD-----LEKDPFAEVHHPGDLKNSVEVALNKL--- 321  
QY 376 QARRKEVTDEIVKEFMTF--RKLS 397  
DB 322 -----DPIREKENTPALKKLA 337

RESULT 8  
US-09-123-615-4  
Sequence 4, Application US/09123615  
Patent No. 6090377  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Murty, Lynn E.  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/123,615  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/705,868  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0117 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1184699  
US-09-123-615-4

Query Match 6.3%; Score 132.5; DB 3; Length 388;  
Best Local Similarity 20.8%; Pred. No. 3.6e-06;  
Matches 80; Conservative 58; Mismatches 121; Indels 125; Gaps 20;

QY 91 GRGPS-SEAMHVGHLIPFTKMLQDVFNPLVIGMTDE--KYLKDLTLDO--AYGD 144  
DB 2 GDAPSPEEKH-----LITRNLOEVLGEKELIKERELKIYWGATTKGPHVAFV 54  
QY 145 AVENAKDIIACGFDINKTFIFSDL-----DYMGSFGFYKNVYKIOKH--YTF 190  
DB 55 PMSKIADFLKAGEV--TILFADLHAYLDNMKAPWELERVSYENVIKAMLESIGVPL 112  
QY 191 NOVKGIFG-----FTDSD-----CIGKISFPAIOA 216  
DB 113 EKLKFKIGTDYOLSKSEYTLDVYRLSSVTOHDSKKAAGAVVQVHEHPILISGLLYPGLQ- 171  
QY 217 PSFNSFPQIFRDRTDIOCLIPCAIDODPYFRMTDVAPRIGYPPKALLHSTFFPALOGA 276  
DB 172 --LDEEYLVK-----DAQF--GGIDQKIFTFPAEKYLPALGYSKRVHLMNPMVGLTG- 220  
QY 277 QTKMSASDPNSSIFLTDPAKQIKTKVKNKHAFA-----SGGRDTIEH----- 317  
DB 221 -SKMSSSEESKIDLLDRKEDYKKLKK-AFCBPGVNNNGVLSPIKHYLPPLKSEFVIL 278  
QY 318 --RQFGNCDVDVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALIEVLOPLIAEH 375  
DB 279 RDEKMGKN-----KTYTAYVD-----LEKDPFAEVHHPGDLKNSVEVALNKL--- 321  
QY 376 QARRKEVTDEIVKEFMTF--RKLS 397  
DB 322 -----DPIREKENTPALKKLA 337

RESULT 9  
US-08-855-910-11  
Sequence 11, Application US/08855910  
Patent No. 6221640  
GENERAL INFORMATION:  
APPLICANT: Tao, Jianshi  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Avruich, Anthony S.  
APPLICANT: Yu, Russell V.  
APPLICANT: Mair, Shanjia  
TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Millitia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/855,910  
 FILING DATE: 14-MAY-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook, David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-08  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781) 861-6240  
 TELEFAX: (781) 861-9540  
 INFORMATION FOR SEQ ID NO: 11:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 418 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-855-910-11

Query Match 5.4%; Score 113.5; DB 4; Length 418;  
 Best Local Similarity 20.5%; Pred. No. 0.00056;  
 Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;

QY 88 LYYGRGSSAMAVGHILPIFTKWLQDVNPVLT----- 123  
 DB 33 LYCGVDPDGTGSMHIGLIPFMMKRFLAGHHPYLLGGGTGTIGDSGRTTERVLTQME 92  
 QY 124 -----OWTDEKYLW-KULTDQAYGDAVEN-----AKDIIA 154  
 DB 93 AVQHNVDLSNOMKLEFGKDAEVTMANNYDWSLSLIDPLRDYGKFNVTMLAKDIVA 152  
 QY 155 CGPDINKTIFPSDLDYMGMS-GEYKAVNVKIQKHTVNOVG----- 195  
 DB 153 S--RLSEGISPTETTYQILQSIDFY---TLHKGNILQIGADQWGNITAGDLIRKE 207  
 QY 196 ----TGF-----DSDICGISFPALQAAFSNSFP--QIFRDRDIOCLICAI 241  
 DB 208 GPEAKVGLTIPMLKADGTRFGKTAGAIWLDPKKTSPEFYQFWLNQD----- 258  
 QY 242 DQDPYFMTSDVAPRIGYKPKALHSTFPPALQAGOTKMSASDPNSIFLDTAKOIKTK 301  
 DB 259 -----RDV---IKYL-----FTFLDKER-----IDALAEKYEKE 286  
 QY 302 VNKIAPSGGDTIEHRQFGNCDVDVSFWYLFLEDDDKLBOIRKDYTGAMLTGELK 361  
 DB 287 PKR--BAQRHAE-----VTRFVHDDALBEAOK--ISEALFSGNIK 326  
 QY 362 KALIEVLQ-----PLAEHQAARKEVTDIEVKEFMTPRK 395  
 DB 327 DLTIEIEIGLEHVFV-EITKAKNIVDMLVDTIEPSK 365

RESULT 10  
 US-09-352-990-28  
 ; Sequence 28, Application US/09352990  
 ; Patent No. 6255090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Famodu, Layo O.  
 ; APPLICANT: Orotco, Buddy  
 ; APPLICANT: Rafalski, Antoni  
 ; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
 ; FILE REFERENCE: BB-1191

CURRENT APPLICATION NUMBER: US/09/352,990  
 ; CURRENT FILING DATE: 1999-07-14  
 ; EARLIER APPLICATION NUMBER: 60/092,866  
 ; EARLIER FILING DATE: July 15, 1998  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 28  
 ; LENGTH: 377  
 ; TYPE: PRT  
 ; ORGANISM: Synecchocystis sp.  
 US-09-352-990-28

Query Match 5.2%; Score 109; DB 4; Length 377;  
 Best Local Similarity 21.0%; Pred. No. 0.0015;  
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;

QY 68 FSHRDMNOVLDAVENKKPFLYLTGSGPSSAMAVGHILPIFTKWLQDVNPVLTQMD 127  
 DB 33 FAHRSTTAM-----DKPRIL-SGVQPTGN-LHLGNVLGAI-RSWVEQ-----QOHY 75  
 QY 128 DEKYLKDL-----TLQAVGDAVENAKDIIAGPDINKTPIF----- 165  
 DB 76 DNEFCVVDLHAITVPHNPQLAQ---DTLTALYIACGIDIOYSTIFVQSHVAHSELA 132  
 QY 166 -----SDLYMGSSGFYKAVNVKIQKHTVNOVGKIFGPTDSDICGISFPALQAPSF 219  
 DB 133 WLNCVTPLNLMLERMIOFKEKAVKQGENVS-----VGLLDYPLMAA----- 174  
 QY 220 SNSFPQIFRDRDIOCL-----IPCAIDQPIFRMTDV-----ARIGPKPA 263  
 DB 175 -----DILLYDADKVPVGEDOKOHELTRDIVIRINDKFGREDAPVLTLPBPL 222  
 QY 264 L-LHSTFPALQAGOTKMSASDPN--SIFLDTAKOIKTKVKNKHAFGSGRDTIEHRQF 320  
 DB 223 IRKEGARVMSLADGTGKMSKDSSELSRINLDPREMIKKTKK----- 266  
 QY 321 GGNCDVDVSFWYLFLEDDDKLEQIRKDYTGAMLTGELKAL----- 364  
 DB 267 ---CKTDPO-RGLWF---DDPERECHNLLTYTLISNOTKEAVQECAMGWCQPKPL 319  
 QY 365 -----IEVLQPLAEHQAARKEVTDIEVYK 388  
 DB 320 TETAIALLEPQAKVAEILLADRGEL-DRITQ 349

RESULT 11  
 US-08-923-867-2  
 ; Sequence 2, Application US/08923867  
 ; Patent No. 5851809  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lawlor, Elizabeth  
 ; TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA SYNTHETASE  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/923,867  
 ; FILING DATE: 04-SEP-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 9619072.3  
 ; FILING DATE: 12-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:

NAME: Gimm1, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-867-2

Query Match 5.1%; Score 108; DB 2; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00067;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

Qy 237 IECALDDPYFRMTD-----VAPRIQYPRPALHSTFPALOGAQTMSA 282  
Db 5 VEVGTQKPMIEQTEIRIVSFNNAYNCVAVPEEGIYENE--RAGRLPGLDG-NAKMSK 61

Qy 283 SDPNSSIFLTDTAKIKTKYKNAHAFSGADTIEHRQFGNCVDVVSFMYLTFF--LEDD 340  
Db 62 S-LNNGIYIADADADTLRKVKMSMYTDPPIHREVDGKIEGN---MVFIYLDVFGRPEDA 116

Qy 341 DLEQIRKDYTSGAMLTGELKKALIEVLOPLIAEQARKKEVTDEI 386  
Db 117 QEIAMKERYQRGIGDVTKRYLLEILRELGPIRERRIEPAKDM 162

RESULT 12  
US-08-928-100-4  
Sequence 4, Application US/08928100  
Patent No. 6046174  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danile  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6046174el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm1, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-4

Query Match 5.1%; Score 108; DB 3; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00067;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

Qy 237 IECALDDPYFRMTD-----VAPRIQYPRPALHSTFPALOGAQTMSA 282  
Db 5 VEVGTQKPMIEQTEIRIVSFNNAYNCVAVPEEGIYENE--RAGRLPGLDG-NAKMSK 61

Qy 283 SDPNSSIFLTDTAKIKTKYKNAHAFSGADTIEHRQFGNCVDVVSFMYLTFF--LEDD 340  
Db 62 S-LNNGIYIADADADTLRKVKMSMYTDPPIHREVDGKIEGN---MVFIYLDVFGRPEDA 116

Qy 341 DLEQIRKDYTSGAMLTGELKKALIEVLOPLIAEQARKKEVTDEI 386  
Db 117 QEIAMKERYQRGIGDVTKRYLLEILRELGPIRERRIEPAKDM 162

RESULT 13  
US-09-183-134-2  
Sequence 2, Application US/09183134  
Patent No. 6165759  
GENERAL INFORMATION:  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,134  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/923,867  
FILING DATE: 04-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm1, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-183-134-2

Query Match 5.1%; Score 108; DB 4; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00067;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPDALHSTFPALOGAQTMSA 282
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVLYVEPEGIYENB--RAGRLPGLDG--NAKMSK 61
QY 283 SDPNSSIFLTDTAQKQIKTKYKNAFSGGRDTIEHRQFGNCVDVVSFMYLTFF--LEDD 340
DB 62 S-INNGIYIADDDADTLRKKYMSWYTDPDHIRVEDPGKIEGN---MVFFHLDVFGRPEDA 116
QY 341 DKLEQIRKDYTSQAMLTGELKKALIEVLQPLIAEHQARRKEVYDEI 386
DB 117 QEIADMKERYQRGGLDVTKRYLLEILEREIGPIRERIRIIFPAKDM 162

RESULT 14
US-09-492-581-4
; Sequence 4, Application US/09492581
; Patent No. 6346409
; GENERAL INFORMATION:
; APPLICANT: Gentry, Danile
; APPLICANT: Greenwood, Claire
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6346409e1 trps
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/492,581
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,100
; FILING DATE: 12-SEP-1997
; APPLICATION NUMBER: 9619072.3
; FILING DATE: 12-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31624-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-492-581-4

Query Match 5.1%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPDALHSTFPALOGAQTMSA 282
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVLYVEPEGIYENB--RAGRLPGLDG--NAKMSK 61
QY 283 SDPNSSIFLTDTAQKQIKTKYKNAFSGGRDTIEHRQFGNCVDVVSFMYLTFF--LEDD 340
DB 62 S-INNGIYIADDDADTLRKKYMSWYTDPDHIRVEDPGKIEGN---MVFFHLDVFGRPEDA 116

```

```

QY 341 DKLEQIRKDYTSQAMLTGELKKALIEVLQPLIAEHQARRKEVYDEI 386
DB 117 QEIADMKERYQRGGLDVTKRYLLEILEREIGPIRERIRIIFPAKDM 162

RESULT 15
US-09-425-666-4
; Sequence 4, Application US/09425666
; Patent No. 6416976
; GENERAL INFORMATION:
; APPLICANT: Gentry, Danile
; APPLICANT: Greenwood, Claire
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6416976e1 trps
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/928,100
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31624-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-425-666-4

Query Match 5.1%; Score 108; DB 4; Length 197;
Best Local Similarity 25.3%; Pred. No. 0.00067;
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPDALHSTFPALOGAQTMSA 282
DB 5 VPGVTDQKPMIEQTRIVRSFNNAVNCVLYVEPEGIYENB--RAGRLPGLDG--NAKMSK 61
QY 283 SDPNSSIFLTDTAQKQIKTKYKNAFSGGRDTIEHRQFGNCVDVVSFMYLTFF--LEDD 340
DB 62 S-INNGIYIADDDADTLRKKYMSWYTDPDHIRVEDPGKIEGN---MVFFHLDVFGRPEDA 116
QY 341 DKLEQIRKDYTSQAMLTGELKKALIEVLQPLIAEHQARRKEVYDEI 386
DB 117 QEIADMKERYQRGGLDVTKRYLLEILEREIGPIRERIRIIFPAKDM 162

Search completed: July 10, 2003, 12:32:51
Job time : 13.1515 sec

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2003, 12:24:54 ; Search time 32.08 Seconds  
(without alignments)  
1665.633 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPMTV.....VTDEIVKEFMTPRKLSDFQ 401

Scoring table: BLOSUM62

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

1. /SID2/gcgdata/genseq/genseqp-emb1/AA1980.DAT.\*  
A\_Genseq.101002.\*  
2. /SID2/gcgdata/genseq/genseqp-emb1/AA1981.DAT.\*  
3. /SID2/gcgdata/genseq/genseqp-emb1/AA1982.DAT.\*  
4. /SID2/gcgdata/genseq/genseqp-emb1/AA1983.DAT.\*  
5. /SID2/gcgdata/genseq/genseqp-emb1/AA1984.DAT.\*  
6. /SID2/gcgdata/genseq/genseqp-emb1/AA1985.DAT.\*  
7. /SID2/gcgdata/genseq/genseqp-emb1/AA1986.DAT.\*  
8. /SID2/gcgdata/genseq/genseqp-emb1/AA1987.DAT.\*  
9. /SID2/gcgdata/genseq/genseqp-emb1/AA1989.DAT.\*  
10. /SID2/gcgdata/genseq/genseqp-emb1/AA1990.DAT.\*  
11. /SID2/gcgdata/genseq/genseqp-emb1/AA1991.DAT.\*  
12. /SID2/gcgdata/genseq/genseqp-emb1/AA1992.DAT.\*  
13. /SID2/gcgdata/genseq/genseqp-emb1/AA1993.DAT.\*  
14. /SID2/gcgdata/genseq/genseqp-emb1/AA1994.DAT.\*  
15. /SID2/gcgdata/genseq/genseqp-emb1/AA1995.DAT.\*  
16. /SID2/gcgdata/genseq/genseqp-emb1/AA1996.DAT.\*  
17. /SID2/gcgdata/genseq/genseqp-emb1/AA1997.DAT.\*  
18. /SID2/gcgdata/genseq/genseqp-emb1/AA1998.DAT.\*  
19. /SID2/gcgdata/genseq/genseqp-emb1/AA1999.DAT.\*  
20. /SID2/gcgdata/genseq/genseqp-emb1/AA2000.DAT.\*  
21. /SID2/gcgdata/genseq/genseqp-emb1/AA2001.DAT.\*  
22. /SID2/gcgdata/genseq/genseqp-emb1/AA2002.DAT.\*  
23. /SID2/gcgdata/genseq/genseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2116	100.0	415	22	AAE147617	Human supermin1 TR
2	2116	100.0	415	23	AAE134393	Human supermin1 trypsin
3	2116	100.0	437	22	AAE147616	Human mini TryPS.
4	2116	100.0	437	23	AAE134392	Human mini trypsin
5	2116	100.0	484	22	AAE147615	Human full-length
6	2116	100.0	484	23	AAE134391	Human-trypsinophany
7	2101	99.3	475	21	AAE582820	Lung cancer associat
8	2096	99.1	471	20	AAE05372	Human HCMV inducib
9	1988	94.0	382	22	AAE147618	Human inactive TryP
10	1988	94.0	392	23	AAE134394	Human inactive try

11	1370.5	64.8	430	22	AB5642621
12	1370.5	64.8	430	22	AB5642621
13	1304	61.6	402	21	AA6236588
14	1304	61.6	426	21	AA6236587
15	1218.5	57.6	424	22	AA6266931
16	968	45.7	192	21	AA6336699
17	831	39.3	173	21	AA6585177
18	808	37.9	385	22	AA6964609
19	328	15.5	85	23	AA6315155
20	226.5	10.7	142	21	AA5682194
21	203.5	9.6	341	23	AA6266949
22	185.5	8.8	341	19	AA6564233
23	185.5	8.8	341	22	AA6478511
24	182.5	8.6	341	22	AA6266955
25	174.5	8.2	341	23	AA6533666
26	164.5	7.8	409	20	AA6542248
27	162.5	7.7	409	20	AA6442447
28	159.5	7.5	379	20	AA6478233
29	158.5	7.2	378	22	AA6566000
30	152	6.9	344	20	AA6354339
31	147	6.9	46	23	AA6351516
32	146	6.9	339	22	AA6558239
33	138.5	6.5	140	21	AA6330322
34	134.5	6.4	525	22	AB6607455
35	134	6.3	372	22	AA6476144
36	134	6.3	372	23	AA6334500
37	132.5	6.3	372	22	AA6476122
38	132.5	6.3	372	23	AA6334688
39	132.5	6.3	536	22	AA6476111
40	132.5	6.3	536	22	AA6334544
41	128	6.0	334	22	AA6545477
42	126.5	6.0	419	23	AB6495417
43	123	5.8	347	21	AA6319044
44	122.5	5.8	346	21	AA6486177
45	119	5.6	415	23	AA6791655

## ALIGNMENTS

	RESULT 1
AA847617	
ID	AA847617 standard; Protein; 415 AA.
AC	AA847617;
XX	
DT	07-JAN-2002 (first entry)
XX	
DE	Human superinini TyPRS.
XX	
KW	Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain
KM	vascular endothelial cell function; burn; plastic surgery; abdomen;
KM	polyomorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KM	angiogenesis; graft; myocardial infarction; solid tumour; wound healing
KM	dermal ulcers; diabetic ulcer; endothelialization;
KM	tyreophanyl-tRNA synthetase; tyPRS; vascular graft surgery.
XX	
OS	Homo sapiens.
XX	
PN	WO200174841-A1.
PD	11-OCT-2001.
XX	
PF	21-MAR-2001; 2001MO-US08966.
XX	
PR	31-MAR-2000; 2000US-193471P.
XX	
PA	(SCRI ) SCRIPPS RES INST.
PI	Schimmel P, Wakaesugi K;
DR	WPI; 2001-626377/72.
DR	N-PSDB; AAAH3604.
OR	

XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating  
 PT angiogenesis, tumor metastasis and treating myocardial infarction  
 PS Disclosure; Page 129-30; 150pp; English.

XX The sequences given in AAB47615-18 show full length and truncated  
 CC versions of tyrophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain; and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery where reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic RNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 22; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWYVQTSASAKIDYDKLIVRGSSKIDKELINRERATGQRP 60  
 DB 2 SNHGPDATAEEDFVDPWYVQTSASAKIDYDKLIVRGSSKIDKELINRERATGQRP 61  
 QY 61 FLRGIFESHRDNNQVLDAYENKPFYLYTGRPSSEAMVGHILPFIPTKMLQDVFNVP 120  
 DB 62 FLRGIFESHRDNNQVLDAYENKPFYLYTGRPSSEAMVGHILPFIPTKMLQDVFNVP 121  
 QY 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 180  
 DB 122 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 181  
 QY 181 VKIKQKAVTNOYKGIKGFSDSCIGKISPPALAAFSFNSFPQIFRDRDIOCLIPCA 240  
 DB 182 VKIKQKAVTNOYKGIKGFSDSCIGKISPPALAAFSFNSFPQIFRDRDIOCLIPCA 241  
 QY 241 IDDDPYFRMRDVAIRIGYKPKPALHSTFPFALOGAQTCKMSASPNSISFILTAKQIKT 300  
 DB 242 IDDDPYFRMRDVAIRIGYKPKPALHSTFPFALOGAQTCKMSASPNSISFILTAKQIKT 301  
 QY 301 KVKKHAESGGRDTEIEHRQFGANGDVVSFWYLTFFLEDDDKLEQIRKDYTSQAMLTGEL 360  
 DB 302 KVKKHAESGGRDTEIEHRQFGANGDVVSFWYLTFFLEDDDKLEQIRKDYTSQAMLTGEL 361  
 QY 361 KKALIEVLOPLIAEHQARKKEVTDIVKEPRTPKLSFDQ 401  
 DB 362 KKALIEVLOPLIAEHQARKKEVTDIVKEPRTPKLSFDQ 402

RESULT 2  
 AAE13493  
 ID AAE13493 standard; Protein; 415 AA.  
 AC AAE13493;  
 DT 12-FEB-2002 (first entry)  
 DE Human supermini tyrophanyl t-RNA synthetase in pET20B.  
 XX Human; tyrophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;

KM TyRS; vascular endothelial cell function; angiogenesis; wound healing;  
 KM re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
 KM diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
 KM angiography; gene therapy; tumour; inflammation; vascular permeability;  
 KM rheumatoid arthritis; psoriasis; diabetic retinopathy.

XX Homo sapiens.  
 OS  
 XX NO200175078-A1.  
 PN  
 XX 11-OCT-2001.  
 PD  
 XX 21-MAR-2001; 2001MO-US08975.  
 PF  
 XX 31-MAR-2000; 2000US-193471P.  
 PR  
 XX (SCRI) SCRIPPS RES INST.  
 PA  
 XX Schimmel P, Makaevgi K;  
 PI  
 XX WPI; 2002-010784/01.  
 DR  
 XX N-PSDB; AAD22484.  
 DR  
 XX Novel truncated tyrophanyl-tRNA synthetase polypeptides capable of  
 PT regulating vascular endothelial cell function, preferably angiogenesis,  
 PT is useful for treating solid tumor or suppressing tumor metastasis in  
 PT mammal -

XX Example 1; Page 129-130; 149pp; English.

PS The patent discloses human aminoacyl tRNA synthetases, particularly  
 XX truncated tyrophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyRS). TrpRS  
 CC sequences are useful for regulating vascular endothelial cell function,  
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
 CC healing agents for re-vascularising damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers), including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
 CC sequences can also be used in plastic surgery when reconstruction is  
 CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TrpRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
 CC immunoassays to detect the presence of tumours. They are also useful  
 CC for blocking endogenous angiogenic activity and retard the growth of  
 CC solid tumours. These antibodies may also be used to treat inflammation  
 CC caused by increased vascular permeability. Inhibiting the activity of  
 CC TrpRS by antisense technology is useful for preventing further growth  
 CC or even regressing solid tumours, and for treating rheumatoid arthritis,  
 CC psoriasis, diabetic retinopathy, all of which are characterised by  
 CC abnormal angiogenesis. The present sequence is human truncated  
 CC tyrophanyl t-RNA synthetase (supermini TrpRS; residues 71-471 of  
 CC full-length TrpRS protein) in pET20B.

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 23; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWYVQTSASAKIDYDKLIVRGSSKIDKELINRERATGQRP 60  
 DB 2 SNHGPDATAEEDFVDPWYVQTSASAKIDYDKLIVRGSSKIDKELINRERATGQRP 61  
 QY 61 FLRGIFESHRDNNQVLDAYENKPFYLYTGRPSSEAMVGHILPFIPTKMLQDVFNVP 120  
 DB 62 FLRGIFESHRDNNQVLDAYENKPFYLYTGRPSSEAMVGHILPFIPTKMLQDVFNVP 121  
 QY 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDMGSSGFYKN 180



```

DB 122 LVIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGPDINKTPIFSDLDYMGSSGFYKN 181
QY 181 VKIKQHTFENVKGIFFGTTSDCIKISFPAIOAAPSPNSPFOIFRDRDIOCLIPCA 240
DB 182 VKIKQHTFENVKGIFFGTTSDCIKISFPAIOAAPSPNSPFOIFRDRDIOCLIPCA 241
QY 241 IDODPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKOIKT 300
DB 242 IDODPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKOIKT 301
QY 301 KVNKAHFGSGRDTIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 302 KVNKAHFGSGRDTIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 361
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401
DB 362 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 402

```

## RESULT 3

AAB47616 standard; Protein; 437 AA.

AAB47616; 07-JAN-2002 (first entry)

Human mini TrpRS.

Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain; vascular endothelial cell function; burn; plastic surgery; abdomen; polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis; dermal ulcer; diabetic ulcer; myocardial infarction; solid tumour; wound healing; tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

Homo sapiens.

WO200174841-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US08966.

31-MAR-2000; 2000US-193471P.

(SCRI) SCRIPPS RES INST.

Schimmel P, Wakaugui K;

WPI; 2001-626377/72.

N-PSDB; AAH43603.

New human truncated tyrosyl-tRNA synthetase polypeptide for regulating vascular endothelial function, in particular for regulating angiogenesis, tumor metastasis and treating myocardial infarction - Disclosure; Page 123-24; 150pp; English.

The sequences given in AAB47615-18 show full length and truncated versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of the invention comprises a Rossmann fold nucleotide binding domain, and is capable of regulating vascular endothelial cell function. It is of approx. 40 kDa dalton molecular weight and is produced by cleavage of full length TrpRS with polymorphonuclear leucocyte elastase. Truncated TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing angiogenesis to a graft, treating myocardial infarction, solid tumor, and a condition that would benefit from increased or decreased angiogenesis in a mammal, in particular humans. It is also useful in diagnosis and as a wound healing agent for treating wounds such as dermal ulcers, diabetic ulcers, burns and injuries and in plastic surgery when reconstruction is required following a burn or

for cosmetic purposes. It is particularly useful in the treatment of abdominal wounds where there is high risk of infection. Truncated TrpRS promotes endothelialization in vascular graft surgery and is used in conjunction with angiography to administer the angiogenic tRNA synthetase polypeptides or polynucleotides directly to the lumen and wall of the blood vessel.

SO Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 22; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.8e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 SNHGDAHAEEDPFDPMTVOTSSAKGIDYDKLIVRFGSSKIDKLINRIEATGQRPKH 60
DB 24 SNHGDAHAEEDPFDPMTVOTSSAKGIDYDKLIVRFGSSKIDKLINRIEATGQRPKH 83
QY 61 FLRRGIFFSHRDMNOVLDAENKPPVLYTGRPSSEAMHGHLPFTTKVLQDVFNVP 120
DB 84 FLRRGIFFSHRDMNOVLDAENKPPVLYTGRPSSEAMHGHLPFTTKVLQDVFNVP 143
QY 121 LVIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGPDINKTPIFSDLDYMGSSGFYKN 180
DB 144 LVIQMTDDEKYLMDLTLDQAYGDAVENAKDIIACGPDINKTPIFSDLDYMGSSGFYKN 203
QY 181 VKIKQHTFENVKGIFFGTTSDCIKISFPAIOAAPSPNSPFOIFRDRDIOCLIPCA 240
DB 204 VKIKQHTFENVKGIFFGTTSDCIKISFPAIOAAPSPNSPFOIFRDRDIOCLIPCA 263
QY 241 IDODPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKOIKT 300
DB 264 IDODPYFMTTRDVARIGYKPKPALHSTFPPALOGAQTMSADPNSSIFLDTAKOIKT 323
QY 301 KVNKAHFGSGRDTIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360
DB 324 KVNKAHFGSGRDTIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 383
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401
DB 384 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 424

```

## RESULT 4

AAE13492 standard; Protein; 437 AA.

AAE13492;

12-FEB-2002 (first entry)

Human mini tryptophanyl t-RNA synthetase in pET20B.

Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl-tRNA synthetase; TrpRS; vascular endothelial cell function; angiogenesis; wound healing; re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury; diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction; angiography; gene therapy; tumour; inflammation; vascular permeability; rheumatoid arthritis; psoriasis; diabetic retinopathy.

Homo sapiens.

WO200175078-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US08975.

31-MAR-2000; 2000US-193471P.

(SCRI) SCRIPPS RES INST.

Schimmel P, Wakaugui K;

XX

DR WPI: 2002-010784/01.  
DR N-PsDB; AAB22483.

PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
PT regulating vascular endothelial cell function, preferably angiogenesis,  
PT is useful for treating solid tumor or suppressing tumor metastasis in  
PT mammal -

XX Example 1; Page 123-124; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly  
XX truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
XX fold nucleotide binding domain and polynucleotides encoding them. The  
XX invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
XX sequences are useful for regulating vascular endothelial cell function,  
XX preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
XX healing agents for re-vascularizing damaged tissues. They are useful for  
XX treating full-thickness wounds (e.g. dermal ulcers), including pressure  
XX sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
XX sequences can also be used in plastic surgery when reconstruction is  
XX required following a burn, other trauma, or even for cosmetic purposes.  
XX Angiogenic TrpRS is also used in association with surgery and following  
XX the repair of cuts, for promoting endothelialisation in vascular graft  
XX surgery and for repairing the damage of myocardial infarction and in  
XX conjunction with coronary bypass surgery by stimulating the growth of  
XX transplanted tissue. TrpRS is also used in conjunction with angiography.  
XX TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
XX immunassays to detect the presence of tumours. They are also useful  
XX for blocking endogenous angiogenic activity and retard the growth of  
XX solid tumours. These antibodies may also be used to treat inflammation  
XX caused by increased vascular permeability. Inhibiting the activity of  
XX TrpRS by antisense technology is useful for preventing further growth  
XX or even regress solid tumours, and for treating rheumatoid arthritis,  
XX psoriasis, diabetic retinopathy, all of which are characterised by  
XX abnormal angiogenesis. The present sequence is human truncated  
XX tryptophanyl t-RNA synthetase (mini TrpRS; residues 48-471 of  
XX full-length TrpRS protein) protein in PET20B.

XX Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 23; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.8e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDPVDPMVTQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 60  
DB 24 SNHGPDTEAEEDPVDPMVTQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 83  
QY 61 FLRGGIFFSHRDNNQVLDAVENKKPFLYTGSGPSSSEAMVGHILPFIPTKWLQDVFNVP 120  
DB 84 FLRGGIFFSHRDNNQVLDAVENKKPFLYTGSGPSSSEAMVGHILPFIPTKWLQDVFNVP 143  
QY 121 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSGLDYMGSSEGYKN 180  
DB 144 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSGLDYMGSSEGYKN 203  
QY 181 VVKTIQKVTNNQVKGIFGFTSDGIGKISPPAIDAAFSFNSPQIRDRDITDQCLIPCA 240  
DB 204 VVKTIQKVTNNQVKGIFGFTSDGIGKISPPAIDAAFSFNSPQIRDRDITDQCLIPCA 263  
QY 241 IDDDPYFRMTRDVAIRIGYRPPALHSTFPALQAGTQKMSASDPNSSIFLDTAKOIKT 300  
DB 264 IDDDPYFRMTRDVAIRIGYRPPALHSTFPALQAGTQKMSASDPNSSIFLDTAKOIKT 323  
QY 301 KVNKHAFSGGRDTEEHRQFGKNCDDVVSFMYLTLPFLIEDDKLEQIKDYTSGAMLTGEL 360  
DB 324 KVNKHAFSGGRDTEEHRQFGKNCDDVVSFMYLTLPFLIEDDKLEQIKDYTSGAMLTGEL 383  
QY 361 KKALILEVLQPLIAEHQARRKEVNDIYEKPEPTPKLSFDFQ 401  
DB 384 KKALILEVLQPLIAEHQARRKEVNDIYEKPEPTPKLSFDFQ 424

RESULT 5

ID AAB47615 standard; Protein; 484 AA.

XX AAB47615;

AC 07-JAN-2002 (first entry)

DE Human full-length TrpRS.

KW Tyrosyl-tRNA synthetase; TyrRS; Rossmann fold nucleotide binding domain;  
KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
KW dermal ulcer; diabetic ulcer; endothelialization;  
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

XX Homo sapiens.

PN W0200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US08966.

PR 31-MAR-2000; 2000US-193471P.

XX (SCRI ) SCRIPPS RES INST.

XX Schimmel P, Wakauegi K;

XX WPI: 2001-626377/72.

DR N-PsDB; AAB43602.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction

XX Disclosure; Page 117-119; 150pp; English.

XX The sequences given in AAB47615-18 show full length and truncated  
XX versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
XX the invention comprises a Rossmann fold nucleotide binding domain, and  
XX is capable of regulating vascular endothelial cell function. It is of  
XX approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
XX full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
XX TrpRS is useful for regulating angiogenesis, tumor metastasis,  
XX enhancing angiogenesis to a graft, treating myocardial infarction,  
XX solid tumor, and a condition that would benefit from increased or  
XX decreased angiogenesis in a mammal, in particular humans. It is also  
XX useful in diagnosis and as a wound healing agent for treating wounds  
XX such as dermal ulcers, diabetic ulcers, burns and injuries and in  
XX plastic surgery when reconstruction is required following a burn or  
XX for cosmetic purposes. It is particularly useful in the treatment of  
XX abdominal wounds where there is high risk of infection. Truncated TrpRS  
XX promotes endothelialization in vascular graft surgery and is used in  
XX conjunction with angiography to administer the angiogenic tRNA  
XX synthetase polypeptides or polynucleotides directly to the lumen and  
XX wall of the blood vessel.

XX Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 2.1e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDPVDPMVTQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 60  
DB 71 SNHGPDTEAEEDPVDPMVTQTSASAKIDYDKLIVFGSSKIDKELINRIERATGQRPH 130  
QY 61 FLRGGIFFSHRDNNQVLDAVENKKPFLYTGSGPSSSEAMVGHILPFIPTKWLQDVFNVP 120  
DB 131 FLRGGIFFSHRDNNQVLDAVENKKPFLYTGSGPSSSEAMVGHILPFIPTKWLQDVFNVP 190

QY 121 LVYQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 180  
 DB 191 LVYQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 250  
 QY 181 VKIKGHVTFNVKGIFFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 240  
 DB 251 VKIKGHVTFNVKGIFFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 310  
 QY 241 IDDPYFMTDVAPRIGYKPKALHSTFPALQAGQTMASADPNSIFLDTAKQIKT 300  
 DB 311 IDDPYFMTDVAPRIGYKPKALHSTFPALQAGQTMASADPNSIFLDTAKQIKT 370  
 QY 301 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 371 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 401  
 DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 471

RESULT 6  
 AAE13491  
 ID AAE13491 standard; Protein; 484 AA.  
 XX AAE13491;  
 AC 12-FEB-2002 (first entry)  
 DT 12-FEB-2002 (first entry)  
 XX  
 DE Human tryptophanyl t-RNA synthetase (TrpRS) in PET20B.  
 XX  
 KM Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;  
 KM TyRS; vascular endothelial cell function; angiogenesis; wound healing;  
 KM re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
 KM diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
 KM angiography; gene therapy; tumour; inflammation; vascular permeability;  
 KM rheumatoid arthritis; psoriasis; diabetic retinopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175078-A1.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US08975.  
 XX  
 PR 31-MAR-2000; 2000US-193471P.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX  
 PI Schimmel P, Wakaugui K;  
 XX  
 DR WPI; 2002-010784/01.  
 DR N-PSDB; AAD22482.  
 XX  
 PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
 PT regulating vascular endothelial cell function, preferably angiogenesis,  
 PT is useful for treating solid tumor or suppressing tumor metastasis in  
 PT mammal  
 XX  
 PS Example 1; Page 117-119; 149pp; English.  
 XX  
 CC The patent discloses human aminoacyl tRNA synthetases, particularly  
 CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rosemann  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyRS). TrpRS  
 CC sequences are useful for regulating vascular endothelial cell function,  
 CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
 CC healing agents for re-vascularizing damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
 CC sequences can also be used in plastic surgery when reconstruction is

CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TrpRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
 CC immunoassays to detect the presence of tumours. They are also useful  
 CC for blocking endogenous angiogenic activity and retard the growth of  
 CC solid tumours. These antibodies may also be used to treat inflammation  
 CC caused by increased vascular permeability. Inhibiting the activity of  
 CC TrpRS by antisense technology is useful for preventing further growth  
 CC or even regress solid tumours, and for treating rheumatoid arthritis,  
 CC psoriasis, diabetic retinopathy, all of which are characterised by  
 CC abnormal angiogenesis. The present sequence is human tryptophanyl  
 CC t-RNA synthetase (TrpRS) in PET20B.  
 XX  
 SQ Sequence 484 AA.  
 Query Match 100.0%; Score 2116; DB 23; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDFYDPWTVQTSAGKIDYDKLIYRFGSSKIDKELINRIERATGQRPNN 60  
 DB 71 SNHGPDATAEEDFYDPWTVQTSAGKIDYDKLIYRFGSSKIDKELINRIERATGQRPNN 130  
 QY 61 FLRRGIFFSHRDMQVLDAYENKKPPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 120  
 DB 131 FLRRGIFFSHRDMQVLDAYENKKPPYLYTGRGSSSEAMHVGHLIPFTKMLQDVFNVP 190  
 QY 121 LVYQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 180  
 DB 191 LVYQMTDEKYLKMDLTDOAGDAVENAKDIIAGGPDINKTIFESDLDMGMSGGFYKN 250  
 QY 181 VKIKGHVTFNVKGIFFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 240  
 DB 251 VKIKGHVTFNVKGIFFGFTDSDCIKISFPALQAPSPNSFPQIFRRTDIQCLIPCA 310  
 QY 241 IDDPYFMTDVAPRIGYKPKALHSTFPALQAGQTMASADPNSIFLDTAKQIKT 300  
 DB 311 IDDPYFMTDVAPRIGYKPKALHSTFPALQAGQTMASADPNSIFLDTAKQIKT 370  
 QY 301 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 371 KVNKAHFGSGRDTIEHRQFGGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 401  
 DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFPDQ 471

RESULT 7  
 AAB58220  
 ID AAB58220 standard; Protein; 475 AA.  
 XX  
 AC AAB58220;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polypeptide sequence SEQ ID 558.  
 XX  
 KM Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KM cardioactive; immunomodulatory; muscular active; vulnary;  
 KM gastroprotective; nephrotoxic; antiinfective; gynecological;  
 KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KM proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055180-A2.  
 XX  
 PD 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US05918.  
 XX PF 12-MAR-1999; 99US-0124270.  
 XX PR (HUMA-) HUMAN GENOME SCI INC.  
 XX PA (ROSE/) ROSEN C A.  
 XX PI Ruben SM;  
 XX WPI; 2000-587514/55.  
 XX DR N-PSDB; AAF18096.  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT anti-gene, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX Claim 11, Page 1052-1053; 1425pp; English.  
 XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytoskeletal; cardiovascular;  
 CC immunomodulatory; muscular; active general; vulnary; gastrointestinal  
 CC general; nephrotoxic; anti-infective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.  
 XX Sequence 475 AA;  
 SQ  
 Query Match 99.3%; Score 2101; DB 21; Length 475;  
 Best Local Similarity 99.5%; Pred. No. 7.1e-208;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDPVDPMVTQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 60  
 DB 75 SNHGPDATAEEDPVDPMVTQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 134  
 QY 61 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVNVP 120  
 DB 135 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVNVP 194  
 QY 121 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFFYN 180  
 DB 195 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFFYN 254  
 QY 181 VVKIQKHTVNOYKGIFFGFTSDPCIKISFPALQAAAPSFSNFPQIFRRTDIOCLIPCA 240  
 DB 255 VVKIQKHTVNOYKGIFFGFTSDPCIKISFPALQAAAPSFSNFPQIFRRTDIOCLIPCA 314  
 QY 241 IDDPYFRMTRDVAIRIGYKPKPALHSTFPALQAGTQKMSASDPNSIFLTDIAKOIKT 300  
 DB 315 IDDPYFRMTRDVAIRIGYKPKPALHSTFPALQAGTQKMSASDPNSIFLTDIAKOIKT 374  
 QY 301 KVNKAHFGSGRDTIEHRQFGNCDVVSFMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 360  
 DB 375 KVNKAHFGSGRDTIEHRQFGNCDVVSFMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 434  
 QY 361 KKALIEVLQPLIAEHOARKREVTDEIVKEFMTPRKLSFDPQ 401  
 DB 435 KKALIEVLQPLIAEHOARKREVTDEIVKEFMTPRKLSFDPQ 475

RESULT 8  
 AAY05372  
 ID AAY05372 standard; Protein; 471 AA.  
 XX AC AAY05372;  
 XX AC AAY05372;  
 DT 30-JUN-1999 (first entry)  
 XX DE Human HCMV inducible gene protein, SEQ ID NO 12.  
 XX KW HCMV inducible gene; cig; human; human cytomagalovirus; interferon;  
 XX KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;  
 XX KM drug screening.  
 XX OS Homo sapiens.  
 XX PN MO913075-A2.  
 XX PD 18-MAR-1999.  
 XX PF 08-SEP-1998; 98MO-US18638.  
 XX PR 22-SEP-1997; 97US-0059725.  
 XX PR 08-SEP-1997; 97US-0058180.  
 XX PA (UYPR-) UNIV PRINCETON.  
 XX PI Cong J, Schenk T, Zhu H;  
 XX WPI; 1999-243729/20.  
 XX DR N-PSDB; AAX33942.  
 XX New isolated human genes  
 PS Claim 3; Page 112-114; 184pp; English.  
 XX This sequence is encoded by a human gene of the invention, and is induced  
 CC to express by both HCMV and interferon (IFN), designated HCMV-inducible  
 CC genes (cig or cigs). The invention also relates to genes that are  
 CC repressed in the presence of HCMV infection, designated HCMV-repressible  
 CC genes (crg or crgs). The products can be used to obtain agents which can  
 CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can  
 CC also be used for the development of drugs that would allow for higher  
 CC dosage IFN treatments without the concomitant toxicity normally  
 CC associated with administering high levels of IFN. The products can also  
 CC be used for detection, diagnosis and drug screening.  
 XX Sequence 471 AA;  
 SQ  
 Query Match 99.1%; Score 2096; DB 20; Length 471;  
 Best Local Similarity 99.3%; Pred. No. 2.3e-207;  
 Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDPVDPMVTQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 60  
 DB 71 SNHGPDATAEEDPVDPMVTQTSASAKGIDYDKLIVFGSSKIDKELINRERATGQRP 130  
 QY 61 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVNVP 120  
 DB 131 FLRRGIFFSHRDMMQVLDAYENKKPFYLYTGRGSSSEAMVGHILPFIPTKMLQDVNVP 190  
 QY 121 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFFYN 180  
 DB 191 LVITQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSLDLYMGSSGFFYN 250  
 QY 181 VVKIQKHTVNOYKGIFFGFTSDPCIKISFPALQAAAPSFSNFPQIFRRTDIOCLIPCA 240  
 DB 251 VVKIQKHTVNOYKGIFFGFTSDPCIKISFPALQAAAPSFSNFPQIFRRTDIOCLIPCA 310  
 QY 241 IDDPYFRMTRDVAIRIGYKPKPALHSTFPALQAGTQKMSASDPNSIFLTDIAKOIKT 300  
 DB 311 IDDPYFRMTRDVAIRIGYKPKPALHSTFPALQAGTQKMSASDPNSIFLTDIAKOIKT 370

Qy	301	KVNKAASGGRTLIEEHROFGAGCDVAFSMYLPFLIEDDDGLAQIRDYSGAMLGEL	360
Db	371	KVNKAASGGRTLIEEHROFGAGCDVAFSMYLPFLIEDDDKLEQIRDYSGAMLGEL	430
Qy	361	KKALIEVLQPLIAEHQARKVEYDIEVKEMTPRKLSPDFQ	401
Db	431	KKALIEVLQPLIAEHQARKREYDIEVKEMTPRKLSPDFQ	471
RESULT 9			
AAAB47618			
XX	AAAB47618	standard; Protein; 392 AA.	
XX	AC		
XX	DT	07-JAN-2002 (first entry)	
XX	XX		
DE	XX	Human inactive TrpRS.	
KM	XX	Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;	
KM	XX	vascular endothelial cell function; burn; plastic surgery; abdomen;	
KM	XX	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;	
KM	XX	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;	
KM	XX	dermal ulcer; diabetic ulcer; endothelialization;	
KM	XX	tyryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.	
OS	XX	Homo sapiens.	
PN	XX	MO200174841-A1.	
PD	XX	11-OCT-2001.	
XX	XX		
PF	XX	21-MAR-2001; 2001MO-US08966.	
PR	XX	31-MAR-2000; 2000US-193471P.	
XX	XX		
PA	XX	(SCRI ) SCRIPPS RES INST.	
XX	XX	Schimmel P, Wakasugi K;	
DR	XX	WPI; 2001-626377/72.	
DR	XX	N-PSDB; AAH43605.	
PT	XX	New human truncated tyrosyl-tRNA synthetase polypeptide for regulating	
PT	XX	vascular endothelial function, in particular for regulating	
XX	XX	angiogenesis, tumor metastasis and treating myocardial infarction -	
PS	XX	Disclosure; Page 135-36; 150pp; English.	
XX	XX		
CC	XX	The sequences given in AAB47615-18 show full length and truncated	
CC	XX	versions of tyryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of	
CC	XX	the invention comprises a Rossmann fold nucleotide binding domain, and	
CC	XX	is capable of regulating vascular endothelial cell function. It is of	
CC	XX	approx. 40 kilo Dalton molecular weight and is produced by cleavage of	
CC	XX	full length TrpRS with polymorphonuclear leucocyte elastase. Truncated	
CC	XX	TrpRS is useful for regulating angiogenesis, tumor metastasis,	
CC	XX	enhancing angiogenesis to a graft, treating myocardial infarction,	
CC	XX	solid tumor, and a condition that would benefit from increased or	
CC	XX	decreased angiogenesis in a mammal, in particular humans. It is also	
CC	XX	useful in diagnosis and as a wound healing agent for treating wounds	
CC	XX	such as dermal ulcers, diabetic ulcers, burns and injuries and in	
CC	XX	plastic surgery when reconstruction is required following a burn or	
CC	XX	for cosmetic purposes. It is particularly useful in the treatment of	
CC	XX	abdominal wounds where there is high risk of infection. Truncated TrpRS	
CC	XX	promotes endothelialization in vascular graft surgery and is used in	
CC	XX	conjunction with angiography to administer the angiogenic tRNA	
CC	XX	synthetase polypeptides or polynucleotides directly to the lumen and	
CC	XX	wall of the blood vessel.	
XX	XX		
Sequence	392 AA;		
Query Match	94.0%;	Score 1988;	DB 22; Length 392;
Best Local Similarity	100.0%;	Pred. No. 2.5e-196;	

	Matches	378;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	
QY		24	SAKGIDYDKLIVRFGSSSKIDKELINRIERATGQRPHHFLRKGIFFSHRDNNVLDAYENK	83						
DB		2	SAKGIDYDKLIVRFSSSKIDKELINRIERATGQRPHHFLRGIFFSHRDNNVLDAYENK	61						
QY		84	KPEYLVTGRGPSEEMHVGHLLPFIETKVLGVFNPLVIQTMDDEKYLMKDITLDOAYG	143						
DB		62	KPEYLVTGRGPSEEMHVGHLLPFIETKVLGVFNPLVIQMTDDKEYLMKDITLDOAYG	121						
QY		144	DAVENAKDIILACGPDINKTFIFSDLDYMGSSSGFYKNVKIKQKVTFENVQKGIFFGTDS	203						
DB		122	DAVENAKDIILACGPDINKTFIFSDLDYMGSSGFYNVVKIKQKVTFENVQKGIFFGTDS	181						
QY		204	CIGKISFPALIQAPSPFNSFPQIFDRRTDIOCLIPCAIDODPYFMTRDPVRIGYPKPA	263						
DB		182	CIGKISFPALIQAPSPFNSFPQIFDRRTDIOCLIPCAIDODPYFMTRDPVRIGYPKPA	241						
QY		264	LHSTFFPALQAOTKMSASDNSSIFLTDTAKOIKTKVNKAFAFGSGRDTEIHRPOGN	323						
DB		242	LHSTFFPALQAOTKMSASDNSSIFLTDTAKOIKTKVNKAFAFGSGRDTEIHRPOGN	301						
QY		324	CDVDSPFWLTFELEDDELQIRKDYTSGLMLTGELEKKALLIEVOPLIAEHQARRKEVT	383						
DB		302	CDVDSPFWLTFELEDDELQIRKDYTSGLMLTGELEKKALLIEVOPLIAEHQARRKEVT	361						
QY		384	DEIVKEFMTPRKLSPDFQ 401							
DB		362	DEIVKEFMTPRKLSPDFQ 379							
Db										
RESULT 10										
AAB13494										
ID		AAB13494	standard; Protein; 392 AA.							
XX										
AC		AAB13494;								
XX										
DT		12-FEB-2002	(first entry)							
XX										
DE			Human inactive tryptophanyl t-RNA synthetase in pET20B.							
XX										
KW			Human; [tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;							
KM			TyrRS; vasacular endothelial cell function; angiogenesis; wound healing;							
KW			re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;							
KM			diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;							
KW			angiography; gene therapy; tumour; inflammation; vascular permeability;							
KX			rheumatoid arthritis; psoriasis; diabetic retinopathy.							
OS			Homo sapiens.							
XX										
PN		WO200175078-A1.								
XX										
PD		11-OCT-2001.								
XX										
Pf		21-MAR-2001; 2001WO-US08975.								
XX										
PR		31-MAR-2000; 2000US-193471P.								
XX										
PA		(SCRI ) SCRIPS RES INST.								
PI		Schimmel P, Wakasugi K;								
XX										
DR		WPI; 2002-010784/O1.								
XX		N-Psdb; AAD22485.								
PT			Novel truncated [tryptophanyl-tRNA synthetase polypeptides capable of							
PS			regulating vasacular endothelial cell function, preferably angiogenesis,							
CC			is useful for treating solid tumor or suppressing tumor metastasis in							
			mammal							
			Disclosure: Page 135-136; 149pp; English.							
			The patent discloses humam aminoacyl tRNA synthetases, particularly							

truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann fold nucleotide binding domain and polynucleotides encoding them. The invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS sequences are useful for regulating vascular endothelial cell function, preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound healing agents for re-vascularizing damaged tissues. They are useful for treating full-thickness wounds (e.g. dermal ulcers, including pressure sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS sequences can also be used in plastic surgery when reconstruction is required following a burn, other trauma, or even for cosmetic purposes. Angiogenic TrpRS is also used in association with surgery and following the repair of cuts, for promoting endothelialisation in vascular graft surgery and for repairing the damage of myocardial infarction and in conjunction with coronary bypass surgery by stimulating the growth of transplanted tissue. TrpRS is also used in conjunction with angiography. TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in immunoassays to detect the presence of tumours. They are also useful for blocking endogenous angiogenic activity and retard the growth of solid tumours. These antibodies may also be used to treat inflammation caused by increased vascular permeability. Inhibiting the activity of TrpRS by antisense technology is useful for preventing further growth or even regressing solid tumours, and for treating rheumatoid arthritis, psoriasis, diabetic retinopathy, all of which are characterised by abnormal angiogenesis. The present sequence is human inactive tryptophanyl t-RNA synthetase (TrpRS) in pET20B.

Sequence 392 AA:

Query Match 94.0%; Score 1988; DB 23; Length 392;  
Best Local Similarity 100.0%; Pred. No. 2.5e-196;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

24 SAKGIDYDKLIVRGSSKIKELINLRERAGRPHFIRGRFFSRDMNVDAENK 83  
|||||  
2 SAKGIDYDKLIVRGSSKIDELINLRERAGRPHFIRGRFFSRDMNVDAENK 61  
|||||  
84 KPFLLYTGRRSSSAMVGHILPFIETKWLQDVNVPLVIOMTDEKYLWKDLTLDOAYG 143  
|||||  
62 KPFLLYTGRRSSSAMVGHILPFIETKWLQDVNVPLVIOMTDEKYLWKDLTLDOAYG 121  
|||||  
144 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIOKHTVNOYKGIFFPTSD 203  
|||||  
122 DAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIOKHTVNOYKGIFFPTSD 181  
|||||  
204 CIGKISPPALQAAFSFNSFPQIFRDRDIOCLPCALIDDPYRMRDVAAPRIGYKPA 263  
|||||  
182 CIGKISPPALQAAFSFNSFPQIFRDRDIOCLPCALIDDPYRMRDVAAPRIGYKPA 241  
|||||  
264 LHSSTFPALQAGTAKMSASDPNSIFLDTAKOIKTKVNHAFSGGRDTEEHROFGCN 323  
|||||  
242 LHSSTFPALQAGTAKMSASDPNSIFLDTAKOIKTKVNHAFSGGRDTEEHROFGCN 301  
|||||  
324 CDVAVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVT 383  
|||||  
302 CDVAVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIEVLQPLIAEHQARRKEVT 361  
|||||  
384 DEIVKEFTPRKLSFDFQ 401  
|||||  
362 DEIVKEFTPRKLSFDFQ 379  
|||||

RESULT 11  
ID ABB64621 standard; Protein; 430 AA.  
XX ABB64621;  
AC ABB64621;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster polypeptide SEQ ID NO 20655.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.

XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001MO-US09231.  
PF  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-0614150.  
PT  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PMD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX N-PSDB; ABL08724.  
PT  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -  
PT  
XX  
XX Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS70272).  
CC  
XX  
XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX  
XX Sequence 430 AA:

Query Match 64.8%; Score 1370.5; DB 22; Length 430;  
Best Local Similarity 64.5%; Pred. No. 1.7e-132;  
Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;

7 ATEAEDFVDPMTVVOSSAKGIDYDKLIVRGSSKIDELINLRERAGRPHFIRGRGI 66  
|||||  
38 ATEAEDFVDPMTVVOSSAKGIDYDKLIVRGSSKIDELINLRERAGRPHFIRGRGI 97  
|||||  
67 FFSHRDMNOVLDAVENKKPPYLYTGRGSSSAMVGHILPFIETKWLQDVNVPLVIOMT 126  
|||||  
98 FFSHRDLHTLTLREGGKPFYLYTGRGSSSAMVGHILPFIETKWLQDVNVPLVIOMT 157  
|||||  
127 DDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIOK 186  
|||||  
158 DDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVVKIOK 217  
|||||  
187 HTFENOVKGIFFGTDSDCIGKISPPALQAAFSFNSFPQIFRDRDIOCLPCALIDDPY 246  
|||||  
218 CTFENOVKGIFFGTDSDCIGKISPPALQAAFSFNSFPQIFRDRDIOCLPCALIDDPY 276  
|||||  
247 FRMTDVAAPRIGYKPAALHSSTFPALQAGTAKMSASDPNSIFLDTAKOIKTKVNHAF 306  
|||||  
277 FRMTDVAAPRIGYKPAALHSSTFPALQAGTAKMSASDPNSIFLDTAKOIKTKVNHAF 336  
|||||  
307 FSGGRDTEEHROFGCNCDVAVSFMYLTFLEDDDKLEQIRKDYTSGLMTGELKKALIE 366  
|||||  
337 FSGGRVSEEHRLGQVPEVDVSYLLKPFLEDDDALEVRVAVSGEMLTGELKKALIE 396  
|||||  
367 VLQPLIAEHQARRKEVTDEIVKEFTPRKLSF 398  
|||||  
397 TLPIVIEQHQAARKLITDEVLDKIFELRPLKF 428  
|||||

RESULT 12

AB67203  
ID AB67203 standard; Protein; 430 AA.  
XX  
AC AB67203;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 26401.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN MO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
XX  
DR WPI; 2001-656660/75.  
XX  
DR N-PSDB; AB111306.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 26401; 21bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (AB57737-AB872072).  
CC  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 430 AA;

Query Match 64.8%; Score 1370.5; DB 22; Length 430;  
Best Local Similarity 64.5%; Fred. No. 1.7e-132;  
Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;

QY 7 ATEAEDFDPMPTVOTSSAKGIDYDKLIVRFGSSKIDKELINIRATGQRPHPFRGI 66  
DB 38 ATRPREDVDPMNVASNDAGVDYDKLIRFGSSKIDKELINIRATGQRPHPFRGI 97  
QY 67 FFSHRMNOVLDAVENKPPYLYTGRSPSEAMVGHLLPFIETKMLQDVNPVLIQMT 126  
DB 98 FFSHRDLHTLTLREQKPPYLYTGRSPSSGLHVLVPIMTKMLQETFDVPLVIQLT 157  
QY 127 DDEKYMKRLTLDOAGDAVENAKDIIAGFOINIKPISDDYMGMSGGFYNNVVKIK 186  
DB 158 DDEKTMKQDKVEDAKIKRENAKDIIAGFOINIKPISDDYMGMSGGFYNNVVKIK 217  
QY 187 HTFNVOKSIGFTDSDCIKISFPALIOAPSSFNPOIFRDRTIOCLIPCAIDODPY 246  
DB 218 CTFNVOKSIGFTGSDDIIGIKISFPALIOAPSSFNPOIFRDRTIOCLIPCAIDODPY 276  
QY 247 FFMTRDVAIRIGYKPKALHSTFPALIOAQYTKMSADPNSSIFLTDIAKQIKTKVAKA 306  
DB 277 FFMTRDVAIRIGYKPKALHSTFPALIOAQYTKMSADPNSSIFLTDIAKQIKTKVAKA 336

QY 307 FSGRDTIEHROFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLWNGELKKALIE 366  
DB 337 FSGRVSVEHKKLGVPEVDVSYQLKFFLEDDAKLEVRVAVSGEMLTSEIKKLAIVE 396  
QY 367 VLQPLAEHQARKKEVTDEIVKEFMTPRKLSF 398  
DB 397 TLTPIVEQHQAARKLITDEVDLKYFELRPLKF 428

RESULT 13  
AAG23698  
ID AAG23698 standard; Protein; 402 AA.  
XX  
AC AAG23698;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27101.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0137882.  
PR 01-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139319.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 25-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148365.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 23-AUG-1999; 99US-0149932.  
 PR 25-AUG-1999; 99US-0150864.  
 PR 25-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 14-OCT-1999; 99US-0159639.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 61.6%; Score 1304; DB 21; Length 402;  
 Best Local Similarity 60.3%; Pred. No. 1,1e-125;  
 Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;  
 Oy 6 DATEAE-EDFVDPWVQTSKAGIDYDKLIYRFGSSKIDKELINRIERATGQRPHEFLR 63



DB 7 DEREASSSEQVNPWEVSAADGGKIDYDKLIDKFGCGLDESLIDRVORLTSPQPHVLR 66  
QY 64 RIIFFSHRDMNOVLDAYENKKPPYLTYTGSPSEAHVGHLLPFIPTKMLQOVENPLVI 123  
DB 67 RSVFFAHRDFNELLDYBERGDKFYLTYGRSPSEALHGLLPFMETKYLQEFKVLVI 126  
QY 124 QMTDDEKYLMDLTLLQOAYGDAVENAKDIIACGFDINKTFFISDLDMGMSGFYGNVVK 183  
DB 127 QLTDEKRSIMKULSVESQRLARENKADIIACGFDYTKTFFISDFPIYV--GAFYGNMVK 184  
QY 184 ICKHVTFNQVKIGFGETSDICIGKISPPALQAPSPSPSPQIFRDRTDIOCLIPCAIDQ 243  
DB 185 VCKCVLTKNMGKIFGSGEDPIAKLSFPPQAVSPSPSPPHLPFGKMLRCLIPCAIDQ 244  
QY 244 DVFRTTRVAPRIGPKPALLHSTFFPALQAGQTMASADPSSIFLITDIAKQITKYN 303  
DB 245 DVFRTTRVAPRIGPKPALLHSTFFPALQAGQTMASADPSSALVYDTSADIKNKIN 304  
QY 304 KHAFIGGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTGAMLTGELTKA 363  
DB 305 RAFIGSGGQSIKHLRGLANLEVDIPVKYLISFLEDDSELEHKKETGEGRMLTGEVKKR 364  
QY 364 LIEVLTQPLAEHQARKEVTDIEVKEFMTPEKLSFDFQ 401  
DB 365 LIEVLTIEVEKRRARAAYTDEWDAFMVAVRPLPSKFE 402

RESULT 14  
AAG23697  
ID AAG23697 standard; Protein; 426 AA.  
AC AAG23697;  
XX  
XX  
DT 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27100.  
DE  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 27100.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridization assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX  
PN EP1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137702.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138547.  
PR 10-JUN-1999; 99US-0138640.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR	26-JUL-1999;	9905-0145216;
PR	26-JUL-1999;	9905-0145217;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146387;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147038;
PR	03-AUG-1999;	9905-0147039;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147925;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	15-AUG-1999;	9905-0149368;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151308;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0153263;
PR	10-SEP-1999;	9905-0153707;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154079;
PR	20-SEP-1999;	9905-0154739;
PR	22-SEP-1999;	9905-0155139;
PR	24-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155539;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156566;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0158293;
PR	13-OCT-1999;	9905-0158294;
PR	13-OCT-1999;	9905-0158295;
PR	14-OCT-1999;	9905-0158329;
PR	14-OCT-1999;	9905-0158330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160980;
PR	22-OCT-1999;	9905-0160981;
PR	22-OCT-1999;	9905-0160989;

PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 28-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match	61.6%;	Score 1304;	DB 21;	Length 426;
Best Local Similarity	60.3%;	Pred. No. 1.2e-125;		
Matches 240;	Conservative 67;	Mismatches 87;	Indels 4;	Gaps 2

QY	6	DATAAB--EDPVDPMVYQOTSAAKIGDVKDLVRFSSSKIDELNIRIRRAGORPHHLR	63
Db	31	DEREABESQVYVNWMEVSASDKGKIDDKLIDKRGCCQLDBSLIDRVORLRSRGHVHLR	90
QY	64	RGIEFFSHRDMNOVLDAAYENKKPFYLYTGRGSSSEAMVGHLLPFIETKMLQDVFNVLVI	123
Db	91	RSVFPAHRDFNEILDAVERGDKFYLYTGRGSSSEALHGHILPFIEMFTKYLQEAFCYPLVI	150
QY	124	QMTDEKYLKMDLTLDDAYGDVAVENAKDIIACGPIINTKPIFSLDLDYMGMSGFYKXNVK	183
Db	151	QLTDEKSIWKNLSEVBESORLARENAKDIACGPIVYTTFFPSFDYVG--GAFYKXNVK	208
QY	184	IQKVYTFNOVYGIQGFMTSDDCIGTISPAIDAAFPSSFOIFPDRNDIOCLIFCAIDQ	243
Db	209	VGKCVTLKAMGIFGFSGEDPIAKLSFFPVAVNVSFPSSFHLPFGKXNKRLICAIQDQ	268
QY	244	DPYFRMTRDVAPIRGYKPRALLHSTFFPALQGAQTKMSASDPNSSIPLTDTAKOIKITXN	303
Db	269	DPYFRMTRDVAPIRGYSKPNALIBSTFFPALQGENKQKASADPNLSALYTTDSAKOIKNKN	328
QY	304	KHAFSGGRDITIEHRQFGKNCDDVVSFMYLTFPLEDDDKLEQIRKDIYSGAMLTEBLKKA	363
Db	329	RYAFSGGQDSTIEKHREIRGANLEVDIIPVYCLISFFLEDDSELHIKKEYGEGRMLTEBVKKR	388
QY	364	LIEVLOPLIAHGOARRKEVTDEIYKEFMTPKLSLSPDQ401	
Db	389	LTEVLTETVEGRRARAAVTDEMDAFMAVBPPLSKFE426	

	RESULT 15
AA066931	
ID	AA066931 standard; Protein; 424 AA
AC	AA066931;
XX	
DT	12-APR-2001 (first entry)
DE	Tryptophanyl-tRNA synthetase.
XX	
FM	Tryptophanyl-tRNA synthetase; enzym
XX	
OS	Candida albicans.
XX	
PN	US6174713-B1.
XX	
PD	16-JAN-2001.
XX	
PF	16-JUN-1997; 97US-0876885.
XX	
PR	16-JUN-1997; 97US-0876885.
XX	
PA	(CUBI-) CUBIST PHARM INC.
XX	
PI	Shen X, Houman F;
XX	
DR	WPI: 2001-201806/20.
DR	N-PSDB; AAF55855.
XX	

XX Shen X, Houman F;  
PI  
XX  
DR WPI; 2001-201806/20  
DR N-PSDB; AAF55855.  
XX

PT New nucleic acids encoding Candida cytoplasmic tryptophanyl-tRNA  
PT (ctRNA) synthetases, useful for producing recombinant ctRNA synthetases  
PT and detecting inhibitor of Candida ctRNA synthetase function  
XX  
XX  
PS Claim 4; Fig 1; 32pp; English.

CC The present sequence is Candida albicans cytoplasmic tryptophanyl-tRNA  
CC (ctRNA) synthetase. ctRNA synthetase is useful for producing recombinant  
CC ctRNA synthetases and detecting inhibitors of ctRNA synthetase.

XX  
SQ Sequence 424 AA;

Query Match 57.6%; Score 1218.5; DB 22; Length 424;  
Best Local Similarity 57.1%; Pred. No. 8.4e-117;  
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 8 TEAEDFDVDPWVQ---TSSAKIDYDYLIVFGSSKIDKELINRIERATGQRPPIHLR 63  
DB 13 TESESEOKITPWEVEGAVVDGKSMGIDYDKLISQFTKHITEETLERFKQVTGEPPHFLK 72  
QY 64 RGIFFSHRMNQVLDAVENKKPEVLYTGRGSSSEAMHVGHLIPITFKWLODVENVPLVI 123  
DB 73 RGVFFSQRDLDRLDLYEHGEPPFLYTGKSPSSDMLGHMVPFLFTKWLQEVFDVPLVI 132  
QY 124 QMTDDEKYLK-DLTLDQAYDAVENAKDIIACGFDINKTFISDDIDYMGSSGFYKNV 182  
DB 133 ELTDEKFLFKQQLITDDYKGAENAKDIIANGFNPENTFISDLQVMG--GAFYENVV 190  
QY 183 KIQKVTENOVKIGFETDSDCIKISPAIQAPSFNSFPQIFRRTDIQCLIPCAID 242  
DB 191 RTSROITSTAKAVFGFTSDDCIGKIHFAISIQIATAFPSSFPDVLGLPPTPCLIPCAID 250  
QY 243 QDPYFRMTDVAPRIGYKPKALLSTFPALOGAOTKMSADPNSSIFLTPAKQIKTKV 302  
DB 251 QDPYFRVCDVDADKLFKTPALHAKFPALOGASTKMSADPTTISIEMGTRAKQIQKKI 310  
QY 303 NGAHFGSGRDITIEHRQFGCNDVDSFMVLTFFLEDDBKLBOIRKDYTSGAMLTGELK 362  
DB 311 NKYAFSGGRATAEHRHELGNPEVDVAFQYLSFYSYDDEKLAQLEGGYRKGEILSGEMK 370  
QY 363 ALIEVLOPLIAHQARKKEVTDIYKEFMTPRKLSF 398  
DB 371 ECITVLOEFVSAYQERRSKVDQVVEKFMKPHKLVF 406

Search completed: July 10, 2003, 12:30:12  
Job time : 33.08 secs

